

**APPLICATION FOR UNITED STATES LETTERS PATENT  
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE  
(Case No. 00,1229)**

Inventors: Christopher J. Paszty  
1027 Westridge Drive  
Ventura, California 93003  
  
Citizen of the United States of America

Jianhua Gong  
2489 Girlington Circle  
Thousand Oaks, California 91362  
  
Citizen of People's Republic of China

Betsy Daugherty  
3091 Stonewood Street  
Simi Valley, California 93062  
  
Citizen of the United States of America

Norma Rogers  
4042 Milano Place  
Moorpark, California 93021  
  
Citizen of the United States of America

Assignee: Amgen, Inc.  
  
A Corporation of the State of Delaware

Title: Leucine-Rich Repeat-Containing G-Protein  
Coupled Receptor-8 Molecules and Uses Thereof

## LEUCINE-RICH REPEAT-CONTAINING G-PROTEIN COUPLED RECEPTOR-8 MOLECULES AND USES THEREOF

5 This application claims the benefit of priority from U.S. Provisional Patent Application No. 60/224,455, filed on August 10, 2000, the disclosure of which is explicitly incorporated by reference herein.

### Field of the Invention

10 The present invention relates to Leucine-Rich Repeat-Containing G-Protein Coupled Receptor-8 (LGR8) polypeptides and nucleic acid molecules encoding the same. The invention also relates to selective binding agents, vectors, host cells, and methods for producing LGR8 polypeptides. The invention further relates to pharmaceutical compositions and methods for the diagnosis, treatment, amelioration, and/or prevention of diseases, disorders, and conditions associated with LGR8 polypeptides.

15

### Background of the Invention

Technical advances in the identification, cloning, expression, and manipulation of nucleic acid molecules and the deciphering of the human genome have greatly accelerated the discovery of novel therapeutics. Rapid nucleic acid sequencing techniques can now generate sequence information at unprecedented rates and, coupled with computational analyses, allow the assembly of overlapping sequences into partial and entire genomes and the identification of polypeptide-encoding regions. A comparison of a predicted amino acid sequence against a database compilation of known amino acid sequences allows one to determine the extent of homology to previously identified sequences and/or structural landmarks. The cloning and expression of a polypeptide-encoding region of a nucleic acid molecule provides a polypeptide product for structural and functional analyses. The manipulation of nucleic acid molecules and encoded polypeptides may confer advantageous properties on a product for use as a therapeutic.

30 In spite of the significant technical advances in genome research over the past decade, the potential for the development of novel therapeutics based on the human

genome is still largely unrealized. Many genes encoding potentially beneficial polypeptide therapeutics or those encoding polypeptides, which may act as "targets" for therapeutic molecules, have still not been identified. Accordingly, it is an object of the invention to identify novel polypeptides, and nucleic acid molecules encoding the same, which have diagnostic or therapeutic benefit.

Members of a glycoprotein hormone receptor subfamily of G-protein coupled receptor (GPCR)/seven-transmembrane domain receptors are characterized by a relatively large (more than 330 amino acids) ligand-binding N-terminal extracellular domain that contains a unique leucine-rich repeat structure (Dufau, 1998, *Annu. Rev. Physiol.* 60:461-96). Among the members of this subfamily are the thyroid stimulating hormone (TSH) receptor, the follicle stimulating hormone (FSH) receptor, and the luteinizing hormone (LH)/chorionic gonadotropin (CG) receptor. Recently, a number of orphan GPCRs having significant homology to the glycoprotein hormone receptor subfamily have been described. These novel members include the leucine-rich repeat-containing G-protein coupled receptor (LGR) 4 (Hsu *et al.*, 1998, *Mol. Endocrinol.* 12:1830-45; PCT Publication No. WO 99/15545), LGR5 (McDonald *et al.*, 1998, *Biochem. Biophys. Res. Commun.* 247:266-70; Hsu *et al.*, 1998, *Mol. Endocrinol.* 12:1830-45; PCT Publication No. WO 99/15660), LGR6 (European Patent App. No. EP 0 950 711 A2), and LGR7 (PCT Publication No. WO 99/48921; Hsu *et al.*, 2000, *Mol. Endocrinol.* 14:1257-71)

The N-terminal extracellular domain of the glycoprotein hormone receptor subfamily retains ligand-binding capability in the absence of both the transmembrane domains and the cytoplasmic C-terminal region. In fact, the N-terminal extracellular domains of the LH, FSH, and TSH receptors, when expressed recombinantly, have been shown to selectively neutralize LH, FSH, or TSH-induced signal transduction (Osuga *et al.*, 1997, *Mol. Endocrinol.* 11:1659-68). In other words, the soluble extracellular domain has been shown to act as a functional antagonist of the signaling pathway of the receptor from which the extracellular domain was derived.

#### Summary of the Invention

The present invention relates to novel LGR8 nucleic acid molecules encoding a polypeptide having significant homology to the glycoprotein hormone receptor subfamily of GPCR/seven-transmembrane domain receptors. This novel member of the glycoprotein hormone receptor subfamily is most closely related to LGR7.

5       The present invention further relates to four distinct LGR8 alternative splicing variants. The LGR8-A coding sequence consists of 18 coding exons that encode a large N-terminal leucine-rich repeat-containing extracellular domain, seven predicted transmembrane domains, and a cytoplasmic C-terminal region. The LGR8-B coding sequence is identical to the LGR8-A coding sequence with the exception that the LGR8-  
10   B coding sequence lacks one of the exons encoding the N-terminal extracellular domain. The LGR8-C coding sequence is identical to the LGR8-A coding sequence with the exception that the LGR8-C coding sequence lacks three of the exons encoding the N-terminal extracellular domain. The LGR8-D coding sequence consists of exons encoding approximately 90% of the N-terminal extracellular domain of the LGR8-B coding  
15   sequence, but lacks exons encoding the transmembrane domains and the cytoplasmic C-terminal region. Thus, LGR8-D is a secreted, N-terminal extracellular domain version of LGR8-B and likely functions as an antagonist of the LGR8 signaling pathway. LGR8-D is truncated very near the C-terminal end of the N-terminal extracellular domain by virtue of the fact that an additional exon, which contains stop codons, is spliced in just 5' of the  
20   exon which encodes the first transmembrane domain of LGR8-A, LGR8-B, and LGR8-C. It is likely that the N-terminal extracellular domains of LGR8-A, LGR8-B, and LGR8-C would be able to function as antagonists of the LGR8 signaling pathway.

25       The invention provides for an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

- (a) the nucleotide sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22;
- (b) a nucleotide sequence encoding the polypeptide as set forth in any of SEQ  
30   ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO:

10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(c) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of either (a) or (b); and

5 (d) a nucleotide sequence complementary to either (a) or (b).

The invention also provides for an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of

(a) a nucleotide sequence encoding a polypeptide which is at least about 70  
10 percent identical to the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID  
15 NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(b) a nucleotide sequence encoding an allelic variant or splice variant of the nucleotide sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6,  
20 SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22, or (a);

(c) a region of the nucleotide sequence of any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22, (a), or (b) encoding a polypeptide fragment of at  
25 least about 25 amino acid residues, wherein the polypeptide fragment has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or is antigenic;

30 (d) a region of the nucleotide sequence of any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16,

SEQ ID NO: 19, or SEQ ID NO: 22, or any of (a) - (c) comprising a fragment of at least about 16 nucleotides;

(e) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of any of (a) - (d); and

5 (f) a nucleotide sequence complementary to any of (a) - (d).

The invention further provides for an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ  
10 ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO:  
10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18,  
SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one conservative  
amino acid substitution, wherein the encoded polypeptide has an activity of the  
polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID  
15 NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO:  
15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO:  
23;

(b) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ  
ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO:  
20 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18,  
SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one amino acid  
insertion, wherein the encoded polypeptide has an activity of the polypeptide set forth in  
any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8,  
SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17,  
25 SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(c) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ  
ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO:  
10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18,  
SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one amino acid  
30 deletion, wherein the encoded polypeptide has an activity of the polypeptide set forth in  
any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8,

SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17,  
SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(d) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ  
ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO:  
5 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18,  
SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 which has a C- and/or N- terminal  
truncation, wherein the encoded polypeptide has an activity of the polypeptide set forth in  
any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8,  
SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17,  
10 SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(e) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ  
ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO:  
10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18,  
SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one modification  
15 selected from the group consisting of amino acid substitutions, amino acid insertions,  
amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the  
encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2,  
SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID  
NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID  
20 NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(f) a nucleotide sequence of any of (a) - (e) comprising a fragment of at least  
about 16 nucleotides;

(g) a nucleotide sequence which hybridizes under moderately or highly  
stringent conditions to the complement of any of (a) - (f); and

25 (h) a nucleotide sequence complementary to any of (a) - (e).

The present invention provides for an isolated polypeptide comprising the amino  
acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ  
ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID  
30 NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID  
NO: 23.

The invention also provides for an isolated polypeptide comprising the amino acid sequence selected from the group consisting of:

5 (a) the amino acid sequence as set forth in any of SEQ ID NO: 3, SEQ ID NO: 8, SEQ ID NO: 13, SEQ ID NO: 18, or SEQ ID NO: 21, optionally further comprising an amino-terminal methionine;

(b) an amino acid sequence for an ortholog of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, 10 SEQ ID NO: 21, or SEQ ID NO: 23;

(c) an amino acid sequence which is at least about 70 percent identical to the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23; 15

(d) a fragment of the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 comprising at least about 25 amino acid residues, wherein the fragment has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or is antigenic; and 20

(e) an amino acid sequence for an allelic variant or splice variant of the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or any of (a) - (c). 25 30



The invention further provides for an isolated polypeptide comprising the amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one conservative amino acid substitution, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(b) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one amino acid insertion, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(c) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one amino acid deletion, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(d) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 which has a C- and/or N- terminal truncation,

wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23; and

- 5 (e) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions,
- 10 C-terminal truncation, and N-terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23.

15

- The invention still further provides for an isolated polypeptide comprising the amino acid sequence as set forth in SEQ ID NO: 2 with at least one conservative amino acid substitution selected from the group consisting of: isoleucine at position 26; valine at position 41; isoleucine at position 55; aspartic acid at position 78; aspartic acid at position 123; arginine at position 130; valine at position 135; methionine at position 142; leucine at position 166; tyrosine at position 167; lysine at position 201; valine at position 204; isoleucine at position 216; glutamic acid at position 217; leucine at position 221; leucine at position 240; leucine at position 252; isoleucine at position 277; methionine at position 288; lysine at position 290; isoleucine at position 324; isoleucine at position 341;
- 20 isoleucine at position 344; aspartic acid at position 350; leucine at position 376; valine at position 420; valine at position 425; valine at position 427; isoleucine at position 434; tyrosine at position 442; arginine at position 444; tyrosine at position 450; isoleucine at position 466; isoleucine at position 471; leucine at position 476; phenylalanine at position 478; glutamic acid at position 481; histidine at position 485; phenylalanine at position
- 25 515; tyrosine at position 521; isoleucine at position 522; tyrosine at position 526; valine at position 531; valine at position 541; isoleucine at position 551; valine at position 552;
- 30

glutamic acid at position 561; phenylalanine at position 562; tyrosine at position 566; tyrosine at position 577; aspartic acid at position 579; isoleucine at position 597; isoleucine at position 603; valine at position 616; isoleucine at position 621; isoleucine at position 626; lysine at position 632; leucine at position 649; isoleucine at position 654; valine at position 675; isoleucine at position 682; glutamic acid at position 700; isoleucine at position 702; tyrosine at position 707; tyrosine at position 709; isoleucine at position 727; valine at position 729; methionine at position 737; methionine at position 745; and leucine at position 749; wherein the polypeptide has an activity of the polypeptide as set forth in SEQ ID NO: 2.

10

Also provided are fusion polypeptides comprising LGR8 amino acid sequences.

The present invention also provides for an expression vector comprising the isolated nucleic acid molecules as set forth herein, recombinant host cells comprising the recombinant nucleic acid molecules as set forth herein, and a method of producing an LGR8 polypeptide comprising culturing the host cells and optionally isolating the polypeptide so produced.

A transgenic non-human animal comprising a nucleic acid molecule encoding an LGR8 polypeptide is also encompassed by the invention. The LGR8 nucleic acid molecules are introduced into the animal in a manner that allows expression and increased levels of an LGR8 polypeptide, which may include increased circulating levels. Alternatively, the LGR8 nucleic acid molecules are introduced into the animal in a manner that prevents expression of endogenous LGR8 polypeptide (*i.e.*, generates a transgenic animal possessing an LGR8 polypeptide gene knockout). The transgenic non-human animal is preferably a mammal, and more preferably a rodent, such as a rat or a mouse.

25

Also provided are derivatives of the LGR8 polypeptides of the present invention.

Additionally provided are selective binding agents such as antibodies and peptides capable of specifically binding the LGR8 polypeptides of the invention. Such antibodies and peptides may be agonistic or antagonistic.

30

Pharmaceutical compositions comprising the nucleotides, polypeptides, or selective binding agents of the invention and one or more pharmaceutically acceptable

formulation agents are also encompassed by the invention. The pharmaceutical compositions are used to provide therapeutically effective amounts of the nucleotides or polypeptides of the present invention. The invention is also directed to methods of using the polypeptides, nucleic acid molecules, and selective binding agents.

5       The LGR8 polypeptides and nucleic acid molecules of the present invention may be used to treat, prevent, ameliorate, and/or detect diseases and disorders, including those recited herein.

10       The present invention also provides a method of assaying test molecules to identify a test molecule that binds to an LGR8 polypeptide. The method comprises contacting an LGR8 polypeptide with a test molecule to determine the extent of binding of the test molecule to the polypeptide. The method further comprises determining whether such test molecules are agonists or antagonists of an LGR8 polypeptide. The present invention further provides a method of testing the impact of molecules on the expression of LGR8 polypeptide or on the activity of LGR8 polypeptide.

15       Methods of regulating expression and modulating (*i.e.*, increasing or decreasing) levels of an LGR8 polypeptide are also encompassed by the invention. One method comprises administering to an animal a nucleic acid molecule encoding an LGR8 polypeptide. In another method, a nucleic acid molecule comprising elements that regulate or modulate the expression of an LGR8 polypeptide may be administered.  
20       Examples of these methods include gene therapy, cell therapy, and anti-sense therapy as further described herein.

25       LGR8 polypeptides can be used for identifying ligands thereof. Various forms of “expression cloning” have been used for cloning ligands for receptors (*See, e.g.*, Davis *et al.*, 1996, *Cell*, 87:1161-69). These and other LGR8 ligand cloning experiments are described in greater detail herein. Isolation of the LGR8 ligand(s) allows for the identification or development of novel agonists and/or antagonists of the LGR8 signaling pathway. Such agonists and antagonists include LGR8 ligand(s), anti-LGR8 ligand antibodies and derivatives thereof, small molecules, or antisense oligonucleotides, any of  
30       those recited herein.

### Brief Description of the Figures

Figures 1A-1D illustrate a nucleotide sequence (SEQ ID NO: 1) encoding human LGR8-A (SEQ ID NO: 2). The predicted signal sequence is indicated (underline);

- 5    Figures 2A-2B illustrate a nucleotide sequence (SEQ ID NO: 4) encoding the N-terminal extracellular domain (absent the signal peptide) of human LGR8-A (SEQ ID NO: 5);

Figures 3A-3D illustrate a nucleotide sequence (SEQ ID NO: 6) encoding human LGR8-B (SEQ ID NO: 7). The predicted signal sequence is indicated (underline);

10

Figures 4A-4B illustrate a nucleotide sequence (SEQ ID NO: 9) encoding the N-terminal extracellular domain (absent the signal peptide) of human LGR8-B (SEQ ID NO: 10);

- 15    Figures 5A-5D illustrate a nucleotide sequence (SEQ ID NO: 11) encoding human LGR8-C (SEQ ID NO: 12). The predicted signal sequence is indicated (underline);

Figures 6A-6B illustrate a nucleotide sequence (SEQ ID NO: 14) encoding the N-terminal extracellular domain (absent the signal peptide) of human LGR8-C (SEQ ID NO: 15);

20

Figures 7A-7B illustrate a nucleotide sequence (SEQ ID NO: 16) encoding human LGR8-D (SEQ ID NO: 17). The predicted signal sequence is indicated (underline);

- 25    Figures 8A-8D illustrate a nucleotide sequence (SEQ ID NO: 19) encoding murine LGR8-A (SEQ ID NO: 20). The predicted signal sequence is indicated (underline);

Figures 9A-9B illustrate a nucleotide sequence (SEQ ID NO: 22) encoding the N-terminal extracellular domain (absent the signal peptide) of murine LGR8-A (SEQ ID NO: 23);

30

Figures 10A-10B illustrate an amino acid sequence alignment of human LGR8-A (upper sequence; SEQ ID NO: 2) and human LGR7 long form (lower sequence; SEQ ID NO: 24);

- 5     Figures 11A-11B illustrate an amino acid sequence alignment of the mature form of human LGR8-A (upper sequence; SEQ ID NO: 3) and the mature form of murine LGR8-A (lower sequence; SEQ ID NO: 21);

- 10     Figure 12 illustrates an amino acid sequence alignment of the N-terminal extracellular domain (absent the signal peptide) of human LGR8-A (upper sequence; SEQ ID NO: 5) and the N-terminal extracellular domain of murine LGR8-A (lower sequence; SEQ ID NO: 23).

#### Detailed Description of the Invention

- 15             The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described. All references cited in this application are expressly incorporated by reference herein.

- 20             It will be appreciated that LGR8-A, LGR8-B, and LGR8-C are membrane-bound polypeptides having an N-terminal extracellular domain, multiple transmembrane domains, and a C-terminal cytoplasmic domain. Accordingly, LGR8-A, LGR8-B, and LGR8-C are useful as targets for agonistic or antagonistic molecules, including, but not limited to, antibodies, fusion polypeptides, carbohydrates, polynucleotides (such as antisense oligonucleotides), or small molecular weight organic molecules.

- 25             Additionally, it will be understood that the N-terminal extracellular domains of LGR8-A, LGR8-B, and LGR8-C can be used as antagonists of the LGR8 signaling pathway, for example, where the N-terminal extracellular domain is fused to an Fc portion of an antibody.

- 30             It will also be appreciated that LGR8-D is a secreted form of the N-terminal extracellular domain of LGR8-B. In this regard LGR8-D may act as an antagonist of the LGR8-B ligand(s). LGR8-D can be used as a target for antagonistic and agonistic molecules, including, but not limited to, antibodies, fusion polypeptides, carbohydrates,

polynucleotides (such as antisense oligonucleotides), or small molecular weight organic molecules. For example, an antagonist specific for LGR8-D would inhibit the antagonistic activity of LGR8-D, thus enhancing the activity of LGR8-D ligand(s) and/or enhancing signaling through LGR8 receptors. Conversely an agonist specific for LGR8-D would enhance the antagonistic activity of LGR8-D, thus diminishing the activity of LGR8-D ligand(s) and/or diminishing signaling through LGR8.

### Definitions

The terms “LGR8 gene” or “LGR8 nucleic acid molecule” or “LGR8 polynucleotide” refer to a nucleic acid molecule comprising or consisting of a nucleotide sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22, a nucleotide sequence encoding the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, and nucleic acid molecules as defined herein.

The term “LGR8 polypeptide allelic variant” refers to one of several possible naturally occurring alternate forms of a gene occupying a given locus on a chromosome of an organism or a population of organisms.

The term “LGR8 polypeptide splice variant” refers to a nucleic acid molecule, usually RNA, which is generated by alternative processing of intron sequences in an RNA transcript of LGR8 polypeptide amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23.

The term “isolated nucleic acid molecule” refers to a nucleic acid molecule of the invention that (1) has been separated from at least about 50 percent of proteins, lipids, carbohydrates, or other materials with which it is naturally found when total nucleic acid is isolated from the source cells, (2) is not linked to all or a portion of a polynucleotide to which the “isolated nucleic acid molecule” is linked in nature, (3) is operably linked to a

polynucleotide which it is not linked to in nature, or (4) does not occur in nature as part of a larger polynucleotide sequence. Preferably, the isolated nucleic acid molecule of the present invention is substantially free from any other contaminating nucleic acid molecule(s) or other contaminants that are found in its natural environment that would interfere with its use in polypeptide production or its therapeutic, diagnostic, prophylactic or research use.

The term "nucleic acid sequence" or "nucleic acid molecule" refers to a DNA or RNA sequence. The term encompasses molecules formed from any of the known base analogs of DNA and RNA such as, but not limited to 4-acetylcytosine, 8-hydroxy-N6-methyladenosine, aziridinyl-cytosine, pseudoisocytosine, 5-(carboxyhydroxymethyl)uracil, 5-fluorouracil, 5-bromouracil, 5-carboxymethylaminomethyl-2-thiouracil, 5-carboxy-methylaminomethyluracil, dihydrouracil, inosine, N6-iso-pentenyladenine, 1-methyladenine, 1-methylpseudouracil, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-methyladenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyamino-methyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarbonyl-methyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, oxybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, N-uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, pseudouracil, queosine, 2-thiocytosine, and 2,6-diaminopurine.

The term "vector" is used to refer to any molecule (*e.g.*, nucleic acid, plasmid, or virus) used to transfer coding information to a host cell.

The term "expression vector" refers to a vector that is suitable for transformation of a host cell and contains nucleic acid sequences that direct and/or control the expression of inserted heterologous nucleic acid sequences. Expression includes, but is not limited to, processes such as transcription, translation, and RNA splicing, if introns are present.

The term "operably linked" is used herein to refer to an arrangement of flanking sequences wherein the flanking sequences so described are configured or assembled so as to perform their usual function. Thus, a flanking sequence operably linked to a coding sequence may be capable of effecting the replication, transcription and/or translation of



the coding sequence. For example, a coding sequence is operably linked to a promoter when the promoter is capable of directing transcription of that coding sequence. A flanking sequence need not be contiguous with the coding sequence, so long as it functions correctly. Thus, for example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

The term "host cell" is used to refer to a cell which has been transformed, or is capable of being transformed with a nucleic acid sequence and then of expressing a selected gene of interest. The term includes the progeny of the parent cell, whether or not the progeny is identical in morphology or in genetic make-up to the original parent, so long as the selected gene is present.

The term "LGR8 polypeptide" refers to a polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 and related polypeptides. Related polypeptides include LGR8 polypeptide fragments, LGR8 polypeptide orthologs, LGR8 polypeptide variants, and LGR8 polypeptide derivatives, which possess at least one activity of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. LGR8 polypeptides may be mature polypeptides, as defined herein, and may or may not have an amino-terminal methionine residue, depending on the method by which they are prepared.

The term "LGR8 polypeptide fragment" refers to a polypeptide that comprises a truncation at the amino-terminus (with or without a leader sequence) and/or a truncation at the carboxyl-terminus of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. The term "LGR8 polypeptide fragment" also refers to amino-terminal and/or carboxyl-terminal truncations of LGR8 polypeptide orthologs, LGR8 polypeptide derivatives, or LGR8 polypeptide variants, or to amino-terminal

and/or carboxyl-terminal truncations of the polypeptides encoded by LGR8 polypeptide allelic variants or LGR8 polypeptide splice variants. LGR8 polypeptide fragments may result from alternative RNA splicing or from *in vivo* protease activity. Membrane-bound forms of an LGR8 polypeptide are also contemplated by the present invention. In preferred embodiments, truncations and/or deletions comprise about 10 amino acids, or about 20 amino acids, or about 50 amino acids, or about 75 amino acids, or about 100 amino acids, or about 200 amino acids, or more than about 200 amino acids. The polypeptide fragments so produced will comprise about 25 contiguous amino acids, or about 50 amino acids, or about 75 amino acids, or about 100 amino acids, or about 150 amino acids, or about 200 amino acids, or more than about 200 amino acids. Such LGR8 polypeptide fragments may optionally comprise an amino-terminal methionine residue. It will be appreciated that such fragments can be used, for example, to generate antibodies to LGR8 polypeptides.

The term "LGR8 polypeptide ortholog" refers to a polypeptide from another species that corresponds to LGR8 polypeptide amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. For example, mouse and human LGR8 polypeptides are considered orthologs of each other.

The term "LGR8 polypeptide variants" refers to LGR8 polypeptides comprising amino acid sequences having one or more amino acid sequence substitutions, deletions (such as internal deletions and/or LGR8 polypeptide fragments), and/or additions (such as internal additions and/or LGR8 fusion polypeptides) as compared to the LGR8 polypeptide amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 (with or without a leader sequence). Variants may be naturally occurring (*e.g.*, LGR8 polypeptide allelic variants, LGR8 polypeptide orthologs, and LGR8 polypeptide splice variants) or artificially constructed. Such LGR8 polypeptide variants may be prepared from the corresponding nucleic acid molecules having a DNA sequence that varies accordingly from the DNA sequence as set forth in

any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22. In preferred embodiments, the variants have from 1 to 3, or from 1 to 5, or from 1 to 10, or from 1 to 15, or from 1 to 20, or from 1 to 25, or from 1 to 50, or from 1 to 75, or from 1 to 100, or more than 100 amino acid substitutions, insertions, additions and/or deletions, wherein the substitutions may be conservative, or non-conservative, or any combination thereof.

The term "LGR8 polypeptide derivatives" refers to the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, LGR8 polypeptide fragments, LGR8 polypeptide orthologs, or LGR8 polypeptide variants, as defined herein, that have been chemically modified. The term "LGR8 polypeptide derivatives" also refers to the polypeptides encoded by LGR8 polypeptide allelic variants or LGR8 polypeptide splice variants, as defined herein, that have been chemically modified.

The term "mature LGR8 polypeptide" refers to an LGR8 polypeptide lacking a leader sequence. A mature LGR8 polypeptide may also include other modifications such as proteolytic processing of the amino-terminus (with or without a leader sequence) and/or the carboxyl-terminus, cleavage of a smaller polypeptide from a larger precursor, N-linked and/or O-linked glycosylation, and the like. Exemplary mature CHL2 polypeptides are depicted by the amino acid sequences of SEQ ID NO: 3, SEQ ID NO: 8, SEQ ID NO: 13, SEQ ID NO: 18, OR SEQ ID NO: 21.

The term "LGR8 fusion polypeptide" refers to a fusion of one or more amino acids (such as a heterologous protein or peptide) at the amino- or carboxyl-terminus of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, LGR8 polypeptide fragments, LGR8 polypeptide orthologs, LGR8 polypeptide variants, or LGR8 derivatives, as defined herein. The term "LGR8 fusion polypeptide" also refers to a fusion of one or more amino acids at the amino- or carboxyl-

terminus of the polypeptide encoded by LGR8 polypeptide allelic variants or LGR8 polypeptide splice variants, as defined herein.

The term “biologically active LGR8 polypeptides” refers to LGR8 polypeptides having at least one activity characteristic of the polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. In addition, an LGR8 polypeptide may be active as an immunogen; that is, the LGR8 polypeptide contains at least one epitope to which antibodies may be raised.

The term “isolated polypeptide” refers to a polypeptide of the present invention that (1) has been separated from at least about 50 percent of polynucleotides, lipids, carbohydrates, or other materials with which it is naturally found when isolated from the source cell, (2) is not linked (by covalent or noncovalent interaction) to all or a portion of a polypeptide to which the “isolated polypeptide” is linked in nature, (3) is operably linked (by covalent or noncovalent interaction) to a polypeptide with which it is not linked in nature, or (4) does not occur in nature. Preferably, the isolated polypeptide is substantially free from any other contaminating polypeptides or other contaminants that are found in its natural environment that would interfere with its therapeutic, diagnostic, prophylactic or research use.

The term “identity,” as known in the art, refers to a relationship between the sequences of two or more polypeptide molecules or two or more nucleic acid molecules, as determined by comparing the sequences. In the art, “identity” also means the degree of sequence relatedness between nucleic acid molecules or polypeptides, as the case may be, as determined by the match between strings of two or more nucleotide or two or more amino acid sequences. “Identity” measures the percent of identical matches between the smaller of two or more sequences with gap alignments (if any) addressed by a particular mathematical model or computer program (*i.e.*, “algorithms”).

The term “similarity” is a related concept, but in contrast to “identity,” “similarity” refers to a measure of relatedness that includes both identical matches and conservative substitution matches. If two polypeptide sequences have, for example, 10/20 identical amino acids, and the remainder are all non-conservative substitutions,

then the percent identity and similarity would both be 50%. If in the same example, there are five more positions where there are conservative substitutions, then the percent identity remains 50%, but the percent similarity would be 75% (15/20). Therefore, in cases where there are conservative substitutions, the percent similarity between two polypeptides will be higher than the percent identity between those two polypeptides.

The term “naturally occurring” or “native” when used in connection with biological materials such as nucleic acid molecules, polypeptides, host cells, and the like, refers to materials which are found in nature and are not manipulated by man. Similarly, “non-naturally occurring” or “non-native” as used herein refers to a material that is not found in nature or that has been structurally modified or synthesized by man.

The terms “effective amount” and “therapeutically effective amount” each refer to the amount of an LGR8 polypeptide or LGR8 nucleic acid molecule used to support an observable level of one or more biological activities of the LGR8 polypeptides as set forth herein.

The term “pharmaceutically acceptable carrier” or “physiologically acceptable carrier” as used herein refers to one or more formulation materials suitable for accomplishing or enhancing the delivery of the LGR8 polypeptide, LGR8 nucleic acid molecule, or LGR8 selective binding agent as a pharmaceutical composition.

The term “antigen” refers to a molecule or a portion of a molecule capable of being bound by a selective binding agent, such as an antibody, and additionally capable of being used in an animal to produce antibodies capable of binding to an epitope of that antigen. An antigen may have one or more epitopes.

The term “selective binding agent” refers to a molecule or molecules having specificity for an LGR8 polypeptide. As used herein, the terms, “specific” and “specificity” refer to the ability of the selective binding agents to bind to human LGR8 polypeptides and not to bind to human non-LGR8 polypeptides. It will be appreciated, however, that the selective binding agents may also bind orthologs of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, that is, interspecies versions thereof, such as mouse and rat LGR8 polypeptides.

The term "transduction" is used to refer to the transfer of genes from one bacterium to another, usually by a phage. "Transduction" also refers to the acquisition and transfer of eukaryotic cellular sequences by retroviruses.

5 The term "transfection" is used to refer to the uptake of foreign or exogenous DNA by a cell, and a cell has been "transfected" when the exogenous DNA has been introduced inside the cell membrane. A number of transfection techniques are well known in the art and are disclosed herein. See, e.g., Graham *et al.*, 1973, *Virology* 52:456; Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratories, 1989); Davis *et al.*, *Basic Methods in Molecular Biology* (Elsevier, 1986);  
10 and Chu *et al.*, 1981, *Gene* 13:197. Such techniques can be used to introduce one or more exogenous DNA moieties into suitable host cells.

The term "transformation" as used herein refers to a change in a cell's genetic characteristics, and a cell has been transformed when it has been modified to contain a new DNA. For example, a cell is transformed where it is genetically modified from its  
15 native state. Following transfection or transduction, the transforming DNA may recombine with that of the cell by physically integrating into a chromosome of the cell, may be maintained transiently as an episomal element without being replicated, or may replicate independently as a plasmid. A cell is considered to have been stably transformed when the DNA is replicated with the division of the cell.

20

#### Relatedness of Nucleic Acid Molecules and/or Polypeptides

It is understood that related nucleic acid molecules include allelic or splice variants of the nucleic acid molecule of any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO:  
25 19, or SEQ ID NO: 22, and include sequences which are complementary to any of the above nucleotide sequences. Related nucleic acid molecules also include a nucleotide sequence encoding a polypeptide comprising or consisting essentially of a substitution, modification, addition and/or deletion of one or more amino acid residues compared to the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5,  
30 SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ

ID NO: 23. Such related LGR8 polypeptides may comprise, for example, an addition and/or a deletion of one or more N-linked or O-linked glycosylation sites or an addition and/or a deletion of one or more cysteine residues.

Related nucleic acid molecules also include fragments of LGR8 nucleic acid  
5 molecules which encode a polypeptide of at least about 25 contiguous amino acids, or about 50 amino acids, or about 75 amino acids, or about 100 amino acids, or about 150 amino acids, or about 200 amino acids, or more than 200 amino acid residues of the LGR8 polypeptide of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15,  
10 SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23.

In addition, related LGR8 nucleic acid molecules also include those molecules which comprise nucleotide sequences which hybridize under moderately or highly stringent conditions as defined herein with the fully complementary sequence of the LGR8 nucleic acid molecule of any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6,  
15 SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22, or of a molecule encoding a polypeptide, which polypeptide comprises the amino acid sequence as shown in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or  
20 SEQ ID NO: 23, or of a nucleic acid fragment as defined herein, or of a nucleic acid fragment encoding a polypeptide as defined herein. Hybridization probes may be prepared using the LGR8 sequences provided herein to screen cDNA, genomic or synthetic DNA libraries for related sequences. Regions of the DNA and/or amino acid sequence of LGR8 polypeptide that exhibit significant identity to known sequences are  
25 readily determined using sequence alignment algorithms as described herein and those regions may be used to design probes for screening.

The term "highly stringent conditions" refers to those conditions that are designed to permit hybridization of DNA strands whose sequences are highly complementary, and to exclude hybridization of significantly mismatched DNAs. Hybridization stringency is  
30 principally determined by temperature, ionic strength, and the concentration of denaturing agents such as formamide. Examples of "highly stringent conditions" for

hybridization and washing are 0.015 M sodium chloride, 0.0015 M sodium citrate at 65-68°C or 0.015 M sodium chloride, 0.0015 M sodium citrate, and 50% formamide at 42°C. See Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual* (2nd ed., Cold Spring Harbor Laboratory, 1989); Anderson *et al.*, *Nucleic Acid Hybridisation: A Practical Approach* Ch. 4 (IRL Press Limited).

More stringent conditions (such as higher temperature, lower ionic strength, higher formamide, or other denaturing agent) may also be used – however, the rate of hybridization will be affected. Other agents may be included in the hybridization and washing buffers for the purpose of reducing non-specific and/or background hybridization. Examples are 0.1% bovine serum albumin, 0.1% polyvinyl-pyrrolidone, 0.1% sodium pyrophosphate, 0.1% sodium dodecylsulfate, NaDodSO<sub>4</sub>, (SDS), ficoll, Denhardt's solution, sonicated salmon sperm DNA (or another non-complementary DNA), and dextran sulfate, although other suitable agents can also be used. The concentration and types of these additives can be changed without substantially affecting the stringency of the hybridization conditions. Hybridization experiments are usually carried out at pH 6.8-7.4; however, at typical ionic strength conditions, the rate of hybridization is nearly independent of pH. See Anderson *et al.*, *Nucleic Acid Hybridisation: A Practical Approach* Ch. 4 (IRL Press Limited).

Factors affecting the stability of DNA duplex include base composition, length, and degree of base pair mismatch. Hybridization conditions can be adjusted by one skilled in the art in order to accommodate these variables and allow DNAs of different sequence relatedness to form hybrids. The melting temperature of a perfectly matched DNA duplex can be estimated by the following equation:

$$T_m(^{\circ}\text{C}) = 81.5 + 16.6(\log[\text{Na}^+]) + 0.41(\% \text{G+C}) - 600/N - 0.72(\% \text{formamide})$$

where N is the length of the duplex formed, [Na<sup>+</sup>] is the molar concentration of the sodium ion in the hybridization or washing solution, %G+C is the percentage of (guanine+cytosine) bases in the hybrid. For imperfectly matched hybrids, the melting temperature is reduced by approximately 1°C for each 1% mismatch.

The term “moderately stringent conditions” refers to conditions under which a DNA duplex with a greater degree of base pair mismatching than could occur under “highly stringent conditions” is able to form. Examples of typical “moderately stringent



conditions” are 0.015 M sodium chloride, 0.0015 M sodium citrate at 50-65°C or 0.015 M sodium chloride, 0.0015 M sodium citrate, and 20% formamide at 37-50°C. By way of example, “moderately stringent conditions” of 50°C in 0.015 M sodium ion will allow about a 21% mismatch.

5 It will be appreciated by those skilled in the art that there is no absolute distinction between “highly stringent conditions” and “moderately stringent conditions.” For example, at 0.015 M sodium ion (no formamide), the melting temperature of perfectly matched long DNA is about 71°C. With a wash at 65°C (at the same ionic strength), this would allow for approximately a 6% mismatch. To capture more distantly  
10 related sequences, one skilled in the art can simply lower the temperature or raise the ionic strength.

A good estimate of the melting temperature in 1M NaCl\* for oligonucleotide probes up to about 20nt is given by:

$$T_m = 2^{\circ}\text{C per A-T base pair} + 4^{\circ}\text{C per G-C base pair}$$

15 \*The sodium ion concentration in 6X salt sodium citrate (SSC) is 1M. See Suggs *et al.*, *Developmental Biology Using Purified Genes* 683 (Brown and Fox, eds., 1981).

High stringency washing conditions for oligonucleotides are usually at a temperature of 0-5°C below the T<sub>m</sub> of the oligonucleotide in 6X SSC, 0.1% SDS.

In another embodiment, related nucleic acid molecules comprise or consist of a  
20 nucleotide sequence that is at least about 70 percent identical to the nucleotide sequence as shown in any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22, or comprise or consist essentially of a nucleotide sequence encoding a polypeptide that is at least about 70 percent identical to the polypeptide as set forth in any of SEQ ID NO: 2,  
25 SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. In preferred embodiments, the nucleotide sequences are about 75 percent, or about 80 percent, or about 85 percent, or about 90 percent, or about 95, 96, 97, 98, or 99 percent identical to the nucleotide sequence as  
30 shown in any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22, or the

nucleotide sequences encode a polypeptide that is about 75 percent, or about 80 percent, or about 85 percent, or about 90 percent, or about 95, 96, 97, 98, or 99 percent identical to the polypeptide sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. Related nucleic acid molecules encode polypeptides possessing at least one activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23.

Differences in the nucleic acid sequence may result in conservative and/or non-conservative modifications of the amino acid sequence relative to the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23.

Conservative modifications to the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 (and the corresponding modifications to the encoding nucleotides) will produce a polypeptide having functional and chemical characteristics similar to those of LGR8 polypeptides. In contrast, substantial modifications in the functional and/or chemical characteristics of LGR8 polypeptides may be accomplished by selecting substitutions in the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 that differ significantly in their effect on maintaining (a) the structure of the molecular backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain.

For example, a "conservative amino acid substitution" may involve a substitution of a native amino acid residue with a nonnative residue such that there is little or no

effect on the polarity or charge of the amino acid residue at that position. Furthermore, any native residue in the polypeptide may also be substituted with alanine, as has been previously described for “alanine scanning mutagenesis.”

Conservative amino acid substitutions also encompass non-naturally occurring amino acid residues that are typically incorporated by chemical peptide synthesis rather than by synthesis in biological systems. These include peptidomimetics, and other reversed or inverted forms of amino acid moieties.

Naturally occurring residues may be divided into classes based on common side chain properties:

- 1) hydrophobic: norleucine, Met, Ala, Val, Leu, Ile;
- 2) neutral hydrophilic: Cys, Ser, Thr;
- 3) acidic: Asp, Glu;
- 4) basic: Asn, Gln, His, Lys, Arg;
- 5) residues that influence chain orientation: Gly, Pro; and
- 6) aromatic: Trp, Tyr, Phe.

For example, non-conservative substitutions may involve the exchange of a member of one of these classes for a member from another class. Such substituted residues may be introduced into regions of the human LGR8 polypeptide that are homologous with non-human LGR8 polypeptides, or into the non-homologous regions of the molecule.

In making such changes, the hydropathic index of amino acids may be considered. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics. The hydropathic indices are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte *et al.*, 1982, *J. Mol. Biol.* 157:105-31). It is known that certain amino acids may be substituted for other amino acids having a similar hydropathic index or score and still retain a similar

biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic indices are within  $\pm 2$  is preferred, those which are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

5 It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity, particularly where the biologically functionally equivalent protein or peptide thereby created is intended for use in immunological embodiments, as in the present case. The greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, *i.e.*, with a biological property of the protein.

10 The following hydrophilicity values have been assigned to these amino acid residues: arginine (+3.0); lysine (+3.0); aspartate ( $+3.0 \pm 1$ ); glutamate ( $+3.0 \pm 1$ ); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline ( $-0.5 \pm 1$ ); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); and tryptophan (-3.4). In making changes based upon similar hydrophilicity values, the substitution of amino acids whose hydrophilicity values are within  $\pm 2$  is preferred, those which are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred. One may also identify epitopes from primary amino acid sequences on the basis of hydrophilicity. These regions are also referred to as "epitopic core regions."

20 Desired amino acid substitutions (whether conservative or non-conservative) can be determined by those skilled in the art at the time such substitutions are desired. For example, amino acid substitutions can be used to identify important residues of the LGR8 polypeptide, or to increase or decrease the affinity of the LGR8 polypeptides described herein. Exemplary amino acid substitutions are set forth in Table I.

Table I

Amino Acid Substitutions

Original Residues	Exemplary Substitutions	Preferred Substitutions
Ala	Val, Leu, Ile	Val

Arg	Lys, Gln, Asn	Lys
Asn	Gln	Gln
Asp	Glu	Glu
Cys	Ser, Ala	Ser
Gln	Asn	Asn
Glu	Asp	Asp
Gly	Pro, Ala	Ala
His	Asn, Gln, Lys, Arg	Arg
Ile	Leu, Val, Met, Ala, Phe, Norleucine	Leu
Leu	Norleucine, Ile, Val, Met, Ala, Phe	Ile
Lys	Arg, 1,4 Diamino-butyric Acid, Gln, Asn	Arg
Met	Leu, Phe, Ile	Leu
Phe	Leu, Val, Ile, Ala, Tyr	Leu
Pro	Ala	Gly
Ser	Thr, Ala, Cys	Thr
Thr	Ser	Ser
Trp	Tyr, Phe	Tyr
Tyr	Trp, Phe, Thr, Ser	Phe
Val	Ile, Met, Leu, Phe, Ala, Norleucine	Leu

5 A skilled artisan will be able to determine suitable variants of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 using well-known techniques. For identifying suitable areas of the molecule that may be changed without destroying biological activity, one skilled in the art may target areas not

believed to be important for activity. For example, when similar polypeptides with similar activities from the same species or from other species are known, one skilled in the art may compare the amino acid sequence of an LGR8 polypeptide to such similar polypeptides. With such a comparison, one can identify residues and portions of the molecules that are conserved among similar polypeptides. It will be appreciated that changes in areas of the LGR8 molecule that are not conserved relative to such similar polypeptides would be less likely to adversely affect the biological activity and/or structure of an LGR8 polypeptide. One skilled in the art would also know that, even in relatively conserved regions, one may substitute chemically similar amino acids for the naturally occurring residues while retaining activity (conservative amino acid residue substitutions). Therefore, even areas that may be important for biological activity or for structure may be subject to conservative amino acid substitutions without destroying the biological activity or without adversely affecting the polypeptide structure.

Additionally, one skilled in the art can review structure-function studies identifying residues in similar polypeptides that are important for activity or structure. In view of such a comparison, one can predict the importance of amino acid residues in an LGR8 polypeptide that correspond to amino acid residues that are important for activity or structure in similar polypeptides. One skilled in the art may opt for chemically similar amino acid substitutions for such predicted important amino acid residues of LGR8 polypeptides.

One skilled in the art can also analyze the three-dimensional structure and amino acid sequence in relation to that structure in similar polypeptides. In view of such information, one skilled in the art may predict the alignment of amino acid residues of LGR8 polypeptide with respect to its three dimensional structure. One skilled in the art may choose not to make radical changes to amino acid residues predicted to be on the surface of the protein, since such residues may be involved in important interactions with other molecules. Moreover, one skilled in the art may generate test variants containing a single amino acid substitution at each amino acid residue. The variants could be screened using activity assays known to those with skill in the art. Such variants could be used to gather information about suitable variants. For example, if one discovered that a change to a particular amino acid residue resulted in destroyed, undesirably reduced, or

unsuitable activity, variants with such a change would be avoided. In other words, based on information gathered from such routine experiments, one skilled in the art can readily determine the amino acids where further substitutions should be avoided either alone or in combination with other mutations.

5       A number of scientific publications have been devoted to the prediction of secondary structure. See Moulton, 1996, *Curr. Opin. Biotechnol.* 7:422-27; Chou *et al.*, 1974, *Biochemistry* 13:222-45; Chou *et al.*, 1974, *Biochemistry* 113:211-22; Chou *et al.*, 1978, *Adv. Enzymol. Relat. Areas Mol. Biol.* 47:45-48; Chou *et al.*, 1978, *Ann. Rev. Biochem.* 47:251-276; and Chou *et al.*, 1979, *Biophys. J.* 26:367-84. Moreover, computer  
10       programs are currently available to assist with predicting secondary structure. One method of predicting secondary structure is based upon homology modeling. For example, two polypeptides or proteins that have a sequence identity of greater than 30%, or similarity greater than 40%, often have similar structural topologies. The recent growth of the protein structural database (PDB) has provided enhanced predictability of  
15       secondary structure, including the potential number of folds within the structure of a polypeptide or protein. See Holm *et al.*, 1999, *Nucleic Acids Res.* 27:244-47. It has been suggested that there are a limited number of folds in a given polypeptide or protein and that once a critical number of structures have been resolved, structural prediction will become dramatically more accurate (Brenner *et al.*, 1997, *Curr. Opin. Struct. Biol.* 7:369-  
20       76).

      Additional methods of predicting secondary structure include "threading" (Jones, 1997, *Curr. Opin. Struct. Biol.* 7:377-87; Sippl *et al.*, 1996, *Structure* 4:15-19), "profile analysis" (Bowie *et al.*, 1991, *Science*, 253:164-70; Gribskov *et al.*, 1990, *Methods Enzymol.* 183:146-59; Gribskov *et al.*, 1987, *Proc. Nat. Acad. Sci. U.S.A.* 84:4355-58),  
25       and "evolutionary linkage" (See Holm *et al.*, *supra*, and Brenner *et al.*, *supra*).

      Preferred LGR8 polypeptide variants include glycosylation variants wherein the number and/or type of glycosylation sites have been altered compared to the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15,  
30       SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. In one embodiment, LGR8 polypeptide variants comprise a greater or a lesser number of

N-linked glycosylation sites than the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. An N-linked glycosylation site is characterized by the sequence: Asn-X-Ser or Asn-X-Thr, wherein the amino acid residue designated as X may be any amino acid residue except proline. The substitution of amino acid residues to create this sequence provides a potential new site for the addition of an N-linked carbohydrate chain. Alternatively, substitutions that eliminate this sequence will remove an existing N-linked carbohydrate chain. Also provided is a rearrangement of N-linked carbohydrate chains wherein one or more N-linked glycosylation sites (typically those that are naturally occurring) are eliminated and one or more new N-linked sites are created. Additional preferred LGR8 variants include cysteine variants, wherein one or more cysteine residues are deleted or substituted with another amino acid (*e.g.*, serine) as compared to the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. Cysteine variants are useful when LGR8 polypeptides must be refolded into a biologically active conformation such as after the isolation of insoluble inclusion bodies. Cysteine variants generally have fewer cysteine residues than the native protein, and typically have an even number to minimize interactions resulting from unpaired cysteines.

In other embodiments, related nucleic acid molecules comprise or consist of a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one amino acid insertion and wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ



ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID  
 NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at  
 least one amino acid deletion and wherein the polypeptide has an activity of the  
 polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID  
 5 NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO:  
 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO:  
 23. Related nucleic acid molecules also comprise or consist of a nucleotide sequence  
 encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID  
 NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO:  
 10 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21,  
 or SEQ ID NO: 23 wherein the polypeptide has a carboxyl- and/or amino-terminal  
 truncation and further wherein the polypeptide has an activity of the polypeptide set forth  
 in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8,  
 SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17,  
 15 SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. Related nucleic  
 acid molecules also comprise or consist of a nucleotide sequence encoding a polypeptide  
 as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7,  
 SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15,  
 SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23  
 20 with at least one modification selected from the group consisting of amino acid  
 substitutions, amino acid insertions, amino acid deletions, carboxyl-terminal truncations,  
 and amino-terminal truncations and wherein the polypeptide has an activity of the  
 polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID  
 NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO:  
 25 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO:  
 23.

In addition, the polypeptide comprising the amino acid sequence of any of SEQ  
 ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO:  
 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18,  
 30 SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or other LGR8 polypeptide, may  
 be fused to a homologous polypeptide to form a homodimer or to a heterologous

polypeptide to form a heterodimer. Heterologous peptides and polypeptides include, but are not limited to: an epitope to allow for the detection and/or isolation of an LGR8 fusion polypeptide; a transmembrane receptor protein or a portion thereof, such as an extracellular domain or a transmembrane and intracellular domain; a ligand or a portion thereof which binds to a transmembrane receptor protein; an enzyme or portion thereof which is catalytically active; a polypeptide or peptide which promotes oligomerization, such as a leucine zipper domain; a polypeptide or peptide which increases stability, such as an immunoglobulin constant region; and a polypeptide which has a therapeutic activity different from the polypeptide comprising the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or other LGR8 polypeptide.

Fusions can be made either at the amino-terminus or at the carboxyl-terminus of the polypeptide comprising the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or other LGR8 polypeptide. Fusions may be direct with no linker or adapter molecule or may be through a linker or adapter molecule. A linker or adapter molecule may be one or more amino acid residues, typically from about 20 to about 50 amino acid residues. A linker or adapter molecule may also be designed with a cleavage site for a DNA restriction endonuclease or for a protease to allow for the separation of the fused moieties. It will be appreciated that once constructed, the fusion polypeptides can be derivatized according to the methods described herein.

In a further embodiment of the invention, the polypeptide comprising the amino acid sequence of ANY OF SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, OR SEQ ID NO: 23, or other LGR8 polypeptide, is fused to one or more domains of an Fc region of human IgG. Antibodies comprise two functionally independent parts, a variable domain known

as “Fab,” that binds an antigen, and a constant domain known as “Fc,” that is involved in effector functions such as complement activation and attack by phagocytic cells. An Fc has a long serum half-life, whereas an Fab is short-lived. Capon *et al.*, 1989, *Nature* 337:525-31. When constructed together with a therapeutic protein, an Fc domain can provide longer half-life or incorporate such functions as Fc receptor binding, protein A binding, complement fixation, and perhaps even placental transfer. *Id.* Table II summarizes the use of certain Fc fusions known in the art.

Table II  
Fc Fusion with Therapeutic Proteins

Form of Fc	Fusion partner	Therapeutic implications	Reference
IgG1	N-terminus of CD30-L	Hodgkin's disease; anaplastic lymphoma; T-cell leukemia	U.S. Patent No. 5,480,981
Murine Fcγ2a	IL-10	anti-inflammatory; transplant rejection	Zheng <i>et al.</i> , 1995, <i>J. Immunol.</i> 154:5590-600
IgG1	TNF receptor	septic shock	Fisher <i>et al.</i> , 1996, <i>N. Engl. J. Med.</i> 334:1697-1702; Van Zee <i>et al.</i> , 1996, <i>J. Immunol.</i> 156:2221-30
IgG, IgA, IgM, or IgE (excluding the first domain)	TNF receptor	inflammation, autoimmune disorders	U.S. Patent No. 5,808,029
IgG1	CD4 receptor	AIDS	Capon <i>et al.</i> , 1989, <i>Nature</i> 337: 525-31
IgG1, IgG3	N-terminus of IL-2	anti-cancer, antiviral	Harvill <i>et al.</i> , 1995, <i>Immunotech.</i> 1:95-105
IgG1	C-terminus of OPG	osteoarthritis; bone density	WO 97/23614
IgG1	N-terminus of leptin	anti-obesity	PCT/US 97/23183, filed December 11, 1997
Human Ig Cγ1	CTLA-4	autoimmune disorders	Linsley, 1991, <i>J. Exp. Med.</i> , 174:561-69

In one example, a human IgG hinge, CH2, and CH3 region may be fused at either the amino-terminus or carboxyl-terminus of the LGR8 polypeptides using methods known to the skilled artisan. In another example, a human IgG hinge, CH2, and CH3

region may be fused at either the amino-terminus or carboxyl-terminus of an LGR8 polypeptide fragment (e.g., the predicted extracellular portion of LGR8 polypeptide).

The resulting LGR8 fusion polypeptide may be purified by use of a Protein A affinity column. Peptides and proteins fused to an Fc region have been found to exhibit a substantially greater half-life *in vivo* than the unfused counterpart. Also, a fusion to an Fc region allows for dimerization/multimerization of the fusion polypeptide. The Fc region may be a naturally occurring Fc region, or may be altered to improve certain qualities, such as therapeutic qualities, circulation time, or reduced aggregation.

Identity and similarity of related nucleic acid molecules and polypeptides are readily calculated by known methods. Such methods include, but are not limited to those described in *Computational Molecular Biology* (A.M. Lesk, ed., Oxford University Press 1988); *Biocomputing: Informatics and Genome Projects* (D.W. Smith, ed., Academic Press 1993); *Computer Analysis of Sequence Data* (Part 1, A.M. Griffin and H.G. Griffin, eds., Humana Press 1994); G. von Heinle, *Sequence Analysis in Molecular Biology* (Academic Press 1987); *Sequence Analysis Primer* (M. Gribskov and J. Devereux, eds., M. Stockton Press 1991); and Carillo *et al.*, 1988, *SIAM J. Applied Math.*, 48:1073.

Preferred methods to determine identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are described in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package, including GAP (Devereux *et al.*, 1984, *Nucleic Acids Res.* 12:387; Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, and FASTA (Altschul *et al.*, 1990, *J. Mol. Biol.* 215:403-10). The BLASTX program is publicly available from the National Center for Biotechnology Information (NCBI) and other sources (Altschul *et al.*, *BLAST Manual* (NCB NLM NIH, Bethesda, MD); Altschul *et al.*, 1990, *supra*). The well-known Smith Waterman algorithm may also be used to determine identity.

Certain alignment schemes for aligning two amino acid sequences may result in the matching of only a short region of the two sequences, and this small aligned region may have very high sequence identity even though there is no significant relationship between the two full-length sequences. Accordingly, in a preferred embodiment, the

selected alignment method (GAP program) will result in an alignment that spans at least 50 contiguous amino acids of the claimed polypeptide.

For example, using the computer algorithm GAP (Genetics Computer Group, University of Wisconsin, Madison, WI), two polypeptides for which the percent sequence identity is to be determined are aligned for optimal matching of their respective amino acids (the “matched span,” as determined by the algorithm). A gap opening penalty (which is calculated as 3X the average diagonal; the “average diagonal” is the average of the diagonal of the comparison matrix being used; the “diagonal” is the score or number assigned to each perfect amino acid match by the particular comparison matrix) and a gap extension penalty (which is usually 0.1X the gap opening penalty), as well as a comparison matrix such as PAM 250 or BLOSUM 62 are used in conjunction with the algorithm. A standard comparison matrix is also used by the algorithm (*see Dayhoff et al., 5 Atlas of Protein Sequence and Structure* (Supp. 3 1978)(PAM250 comparison matrix); Henikoff *et al.*, 1992, *Proc. Natl. Acad. Sci USA* 89:10915-19 (BLOSUM 62 comparison matrix)).

Preferred parameters for polypeptide sequence comparison include the following:

Algorithm: Needleman and Wunsch, 1970, *J. Mol. Biol.* 48:443-53;

Comparison matrix: BLOSUM 62 (Henikoff *et al.*, *supra*);

Gap Penalty: 12

Gap Length Penalty: 4

Threshold of Similarity: 0

The GAP program is useful with the above parameters. The aforementioned parameters are the default parameters for polypeptide comparisons (along with no penalty for end gaps) using the GAP algorithm.

Preferred parameters for nucleic acid molecule sequence comparison include the following:

Algorithm: Needleman and Wunsch, *supra*;

Comparison matrix: matches = +10, mismatch = 0

Gap Penalty: 50

Gap Length Penalty: 3

5 The GAP program is also useful with the above parameters. The aforementioned parameters are the default parameters for nucleic acid molecule comparisons.

Other exemplary algorithms, gap opening penalties, gap extension penalties, comparison matrices, and thresholds of similarity may be used, including those set forth in the Program Manual, Wisconsin Package, Version 9, September, 1997. The particular choices to be made will be apparent to those of skill in the art and will depend on the specific comparison to be made, such as DNA-to-DNA, protein-to-protein, protein-to-DNA; and additionally, whether the comparison is between given pairs of sequences (in which case GAP or BestFit are generally preferred) or between one sequence and a large database of sequences (in which case FASTA or BLASTA are preferred).

#### 15 Nucleic Acid Molecules

The nucleic acid molecules encoding a polypeptide comprising the amino acid sequence of an LGR8 polypeptide can readily be obtained in a variety of ways including, without limitation, chemical synthesis, cDNA or genomic library screening, expression library screening, and/or PCR amplification of cDNA.

20 Recombinant DNA methods used herein are generally those set forth in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1989) and/or *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., Green Publishers Inc. and Wiley and Sons 1994). The invention provides for nucleic acid molecules as described herein and methods for obtaining such molecules.

25 Where a gene encoding the amino acid sequence of an LGR8 polypeptide has been identified from one species, all or a portion of that gene may be used as a probe to identify orthologs or related genes from the same species. The probes or primers may be used to screen cDNA libraries from various tissue sources believed to express the LGR8 polypeptide. In addition, part or all of a nucleic acid molecule having the sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22 may be

30

used to screen a genomic library to identify and isolate a gene encoding the amino acid sequence of an LGR8 polypeptide. Typically, conditions of moderate or high stringency will be employed for screening to minimize the number of false positives obtained from the screening.

5 Nucleic acid molecules encoding the amino acid sequence of LGR8 polypeptides may also be identified by expression cloning which employs the detection of positive clones based upon a property of the expressed protein. Typically, nucleic acid libraries are screened by the binding an antibody or other binding partner (*e.g.*, receptor or ligand) to cloned proteins that are expressed and displayed on a host cell surface. The antibody  
10 or binding partner is modified with a detectable label to identify those cells expressing the desired clone.

Recombinant expression techniques conducted in accordance with the descriptions set forth below may be followed to produce these polynucleotides and to express the encoded polypeptides. For example, by inserting a nucleic acid sequence that  
15 encodes the amino acid sequence of an LGR8 polypeptide into an appropriate vector, one skilled in the art can readily produce large quantities of the desired nucleotide sequence. The sequences can then be used to generate detection probes or amplification primers. Alternatively, a polynucleotide encoding the amino acid sequence of an LGR8 polypeptide can be inserted into an expression vector. By introducing the expression  
20 vector into an appropriate host, the encoded LGR8 polypeptide may be produced in large amounts.

Another method for obtaining a suitable nucleic acid sequence is the polymerase chain reaction (PCR). In this method, cDNA is prepared from poly(A)+RNA or total RNA using the enzyme reverse transcriptase. Two primers, typically complementary to  
25 two separate regions of cDNA encoding the amino acid sequence of an LGR8 polypeptide, are then added to the cDNA along with a polymerase such as *Taq* polymerase, and the polymerase amplifies the cDNA region between the two primers.

Another means of preparing a nucleic acid molecule encoding the amino acid sequence of an LGR8 polypeptide is chemical synthesis using methods well known to the  
30 skilled artisan such as those described by Engels *et al.*, 1989, *Angew. Chem. Intl. Ed.* 28:716-34. These methods include, *inter alia*, the phosphotriester, phosphoramidite, and

H-phosphonate methods for nucleic acid synthesis. A preferred method for such chemical synthesis is polymer-supported synthesis using standard phosphoramidite chemistry. Typically, the DNA encoding the amino acid sequence of an LGR8 polypeptide will be several hundred nucleotides in length. Nucleic acids larger than about 100 nucleotides can be synthesized as several fragments using these methods. The fragments can then be ligated together to form the full-length nucleotide sequence of an LGR8 gene. Usually, the DNA fragment encoding the amino-terminus of the polypeptide will have an ATG, which encodes a methionine residue. This methionine may or may not be present on the mature form of the LGR8 polypeptide, depending on whether the polypeptide produced in the host cell is designed to be secreted from that cell. Other methods known to the skilled artisan may be used as well.

In certain embodiments, nucleic acid variants contain codons which have been altered for optimal expression of an LGR8 polypeptide in a given host cell. Particular codon alterations will depend upon the LGR8 polypeptide and host cell selected for expression. Such "codon optimization" can be carried out by a variety of methods, for example, by selecting codons which are preferred for use in highly expressed genes in a given host cell. Computer algorithms which incorporate codon frequency tables such as "Eco\_high.Cod" for codon preference of highly expressed bacterial genes may be used and are provided by the University of Wisconsin Package Version 9.0 (Genetics Computer Group, Madison, WI). Other useful codon frequency tables include "Celegans\_high.cod," "Celegans\_low.cod," "Drosophila\_high.cod," "Human\_high.cod," "Maize\_high.cod," and "Yeast\_high.cod."

In some cases, it may be desirable to prepare nucleic acid molecules encoding LGR8 polypeptide variants. Nucleic acid molecules encoding variants may be produced using site directed mutagenesis, PCR amplification, or other appropriate methods, where the primer(s) have the desired point mutations (*see Sambrook et al., supra*, and Ausubel *et al., supra*, for descriptions of mutagenesis techniques). Chemical synthesis using methods described by Engels *et al., supra*, may also be used to prepare such variants. Other methods known to the skilled artisan may be used as well.

#### Vectors and Host Cells



A nucleic acid molecule encoding the amino acid sequence of an LGR8 polypeptide is inserted into an appropriate expression vector using standard ligation techniques. The vector is typically selected to be functional in the particular host cell employed (*i.e.*, the vector is compatible with the host cell machinery such that amplification of the gene and/or expression of the gene can occur). A nucleic acid molecule encoding the amino acid sequence of an LGR8 polypeptide may be amplified/expressed in prokaryotic, yeast, insect (baculovirus systems) and/or eukaryotic host cells. Selection of the host cell will depend in part on whether an LGR8 polypeptide is to be post-translationally modified (*e.g.*, glycosylated and/or phosphorylated). If so, yeast, insect, or mammalian host cells are preferable. For a review of expression vectors, see *Meth. Enz.*, vol. 185 (D.V. Goeddel, ed., Academic Press 1990).

Typically, expression vectors used in any of the host cells will contain sequences for plasmid maintenance and for cloning and expression of exogenous nucleotide sequences. Such sequences, collectively referred to as “flanking sequences” in certain embodiments will typically include one or more of the following nucleotide sequences: a promoter, one or more enhancer sequences, an origin of replication, a transcriptional termination sequence, a complete intron sequence containing a donor and acceptor splice site, a sequence encoding a leader sequence for polypeptide secretion, a ribosome binding site, a polyadenylation sequence, a polylinker region for inserting the nucleic acid encoding the polypeptide to be expressed, and a selectable marker element. Each of these sequences is discussed below.

Optionally, the vector may contain a “tag”-encoding sequence, *i.e.*, an oligonucleotide molecule located at the 5’ or 3’ end of the LGR8 polypeptide coding sequence; the oligonucleotide sequence encodes polyHis (such as hexaHis), or another “tag” such as FLAG, HA (hemagglutinin influenza virus), or *myc* for which commercially available antibodies exist. This tag is typically fused to the polypeptide upon expression of the polypeptide, and can serve as a means for affinity purification of the LGR8 polypeptide from the host cell. Affinity purification can be accomplished, for example, by column chromatography using antibodies against the tag as an affinity matrix. Optionally, the tag can subsequently be removed from the purified LGR8 polypeptide by various means such as using certain peptidases for cleavage.

Flanking sequences may be homologous (*i.e.*, from the same species and/or strain as the host cell), heterologous (*i.e.*, from a species other than the host cell species or strain), hybrid (*i.e.*, a combination of flanking sequences from more than one source), or synthetic, or the flanking sequences may be native sequences that normally function to regulate LGR8 polypeptide expression. As such, the source of a flanking sequence may be any prokaryotic or eukaryotic organism, any vertebrate or invertebrate organism, or any plant, provided that the flanking sequence is functional in, and can be activated by, the host cell machinery.

Flanking sequences useful in the vectors of this invention may be obtained by any of several methods well known in the art. Typically, flanking sequences useful herein – other than the LGR8 gene flanking sequences – will have been previously identified by mapping and/or by restriction endonuclease digestion and can thus be isolated from the proper tissue source using the appropriate restriction endonucleases. In some cases, the full nucleotide sequence of a flanking sequence may be known. Here, the flanking sequence may be synthesized using the methods described herein for nucleic acid synthesis or cloning.

Where all or only a portion of the flanking sequence is known, it may be obtained using PCR and/or by screening a genomic library with a suitable oligonucleotide and/or flanking sequence fragment from the same or another species. Where the flanking sequence is not known, a fragment of DNA containing a flanking sequence may be isolated from a larger piece of DNA that may contain, for example, a coding sequence or even another gene or genes. Isolation may be accomplished by restriction endonuclease digestion to produce the proper DNA fragment followed by isolation using agarose gel purification, Qiagen® column chromatography (Chatsworth, CA), or other methods known to the skilled artisan. The selection of suitable enzymes to accomplish this purpose will be readily apparent to one of ordinary skill in the art.

An origin of replication is typically a part of those prokaryotic expression vectors purchased commercially, and the origin aids in the amplification of the vector in a host cell. Amplification of the vector to a certain copy number can, in some cases, be important for the optimal expression of an LGR8 polypeptide. If the vector of choice does not contain an origin of replication site, one may be chemically synthesized based

on a known sequence, and ligated into the vector. For example, the origin of replication from the plasmid pBR322 (New England Biolabs, Beverly, MA) is suitable for most gram-negative bacteria and various origins (*e.g.*, SV40, polyoma, adenovirus, vesicular stomatitis virus (VSV), or papillomaviruses such as HPV or BPV) are useful for cloning  
5 vectors in mammalian cells. Generally, the origin of replication component is not needed for mammalian expression vectors (for example, the SV40 origin is often used only because it contains the early promoter).

A transcription termination sequence is typically located 3' of the end of a polypeptide coding region and serves to terminate transcription. Usually, a transcription  
10 termination sequence in prokaryotic cells is a G-C rich fragment followed by a poly-T sequence. While the sequence is easily cloned from a library or even purchased commercially as part of a vector, it can also be readily synthesized using methods for nucleic acid synthesis such as those described herein.

A selectable marker gene element encodes a protein necessary for the survival  
15 and growth of a host cell grown in a selective culture medium. Typical selection marker genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, tetracycline, or kanamycin for prokaryotic host cells; (b) complement auxotrophic deficiencies of the cell; or (c) supply critical nutrients not available from complex media. Preferred selectable markers are the kanamycin resistance gene, the  
20 ampicillin resistance gene, and the tetracycline resistance gene. A neomycin resistance gene may also be used for selection in prokaryotic and eukaryotic host cells.

Other selection genes may be used to amplify the gene that will be expressed. Amplification is the process wherein genes that are in greater demand for the production of a protein critical for growth are reiterated in tandem within the chromosomes of  
25 successive generations of recombinant cells. Examples of suitable selectable markers for mammalian cells include dihydrofolate reductase (DHFR) and thymidine kinase. The mammalian cell transformants are placed under selection pressure wherein only the transformants are uniquely adapted to survive by virtue of the selection gene present in the vector. Selection pressure is imposed by culturing the transformed cells under  
30 conditions in which the concentration of selection agent in the medium is successively changed, thereby leading to the amplification of both the selection gene and the DNA that

encodes an LGR8 polypeptide. As a result, increased quantities of LGR8 polypeptide are synthesized from the amplified DNA.

A ribosome binding site is usually necessary for translation initiation of mRNA and is characterized by a Shine-Dalgarno sequence (prokaryotes) or a Kozak sequence (eukaryotes). The element is typically located 3' to the promoter and 5' to the coding sequence of an LGR8 polypeptide to be expressed. The Shine-Dalgarno sequence is varied but is typically a polypurine (*i.e.*, having a high A-G content). Many Shine-Dalgarno sequences have been identified, each of which can be readily synthesized using methods set forth herein and used in a prokaryotic vector.

A leader, or signal, sequence may be used to direct an LGR8 polypeptide out of the host cell. Typically, a nucleotide sequence encoding the signal sequence is positioned in the coding region of an LGR8 nucleic acid molecule, or directly at the 5' end of an LGR8 polypeptide coding region. Many signal sequences have been identified, and any of those that are functional in the selected host cell may be used in conjunction with an LGR8 nucleic acid molecule. Therefore, a signal sequence may be homologous (naturally occurring) or heterologous to the LGR8 nucleic acid molecule. Additionally, a signal sequence may be chemically synthesized using methods described herein. In most cases, the secretion of an LGR8 polypeptide from the host cell via the presence of a signal peptide will result in the removal of the signal peptide from the secreted LGR8 polypeptide. The signal sequence may be a component of the vector, or it may be a part of an LGR8 nucleic acid molecule that is inserted into the vector.

Included within the scope of this invention is the use of either a nucleotide sequence encoding a native LGR8 polypeptide signal sequence joined to an LGR8 polypeptide coding region or a nucleotide sequence encoding a heterologous signal sequence joined to an LGR8 polypeptide coding region. The heterologous signal sequence selected should be one that is recognized and processed, *i.e.*, cleaved by a signal peptidase, by the host cell. For prokaryotic host cells that do not recognize and process the native LGR8 polypeptide signal sequence, the signal sequence is substituted by a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, or heat-stable enterotoxin II leaders. For yeast secretion, the native LGR8 polypeptide signal sequence may be substituted by the yeast invertase,

alpha factor, or acid phosphatase leaders. In mammalian cell expression the native signal sequence is satisfactory, although other mammalian signal sequences may be suitable.

In some cases, such as where glycosylation is desired in a eukaryotic host cell expression system, one may manipulate the various presequences to improve glycosylation or yield. For example, one may alter the peptidase cleavage site of a particular signal peptide, or add pro-sequences, which also may affect glycosylation. The final protein product may have, in the -1 position (relative to the first amino acid of the mature protein) one or more additional amino acids incident to expression, which may not have been totally removed. For example, the final protein product may have one or two amino acid residues found in the peptidase cleavage site, attached to the amino-terminus. Alternatively, use of some enzyme cleavage sites may result in a slightly truncated form of the desired LGR8 polypeptide, if the enzyme cuts at such area within the mature polypeptide.

In many cases, transcription of a nucleic acid molecule is increased by the presence of one or more introns in the vector; this is particularly true where a polypeptide is produced in eukaryotic host cells, especially mammalian host cells. The introns used may be naturally occurring within the LGR8 gene especially where the gene used is a full-length genomic sequence or a fragment thereof. Where the intron is not naturally occurring within the gene (as for most cDNAs), the intron may be obtained from another source. The position of the intron with respect to flanking sequences and the LGR8 gene is generally important, as the intron must be transcribed to be effective. Thus, when an LGR8 cDNA molecule is being transcribed, the preferred position for the intron is 3' to the transcription start site and 5' to the poly-A transcription termination sequence. Preferably, the intron or introns will be located on one side or the other (*i.e.*, 5' or 3') of the cDNA such that it does not interrupt the coding sequence. Any intron from any source, including viral, prokaryotic and eukaryotic (plant or animal) organisms, may be used to practice this invention, provided that it is compatible with the host cell into which it is inserted. Also included herein are synthetic introns. Optionally, more than one intron may be used in the vector.

The expression and cloning vectors of the present invention will typically contain a promoter that is recognized by the host organism and operably linked to the molecule

encoding the LGR8 polypeptide. Promoters are untranscribed sequences located upstream (*i.e.*, 5') to the start codon of a structural gene (generally within about 100 to 1000 bp) that control the transcription of the structural gene. Promoters are conventionally grouped into one of two classes: inducible promoters and constitutive promoters. Inducible promoters initiate increased levels of transcription from DNA under their control in response to some change in culture conditions, such as the presence or absence of a nutrient or a change in temperature. Constitutive promoters, on the other hand, initiate continual gene product production; that is, there is little or no control over gene expression. A large number of promoters, recognized by a variety of potential host cells, are well known. A suitable promoter is operably linked to the DNA encoding LGR8 polypeptide by removing the promoter from the source DNA by restriction enzyme digestion and inserting the desired promoter sequence into the vector. The native LGR8 promoter sequence may be used to direct amplification and/or expression of an LGR8 nucleic acid molecule. A heterologous promoter is preferred, however, if it permits greater transcription and higher yields of the expressed protein as compared to the native promoter, and if it is compatible with the host cell system that has been selected for use.

Promoters suitable for use with prokaryotic hosts include the beta-lactamase and lactose promoter systems; alkaline phosphatase; a tryptophan (*trp*) promoter system; and hybrid promoters such as the *tac* promoter. Other known bacterial promoters are also suitable. Their sequences have been published, thereby enabling one skilled in the art to ligate them to the desired DNA sequence, using linkers or adapters as needed to supply any useful restriction sites.

Suitable promoters for use with yeast hosts are also well known in the art. Yeast enhancers are advantageously used with yeast promoters. Suitable promoters for use with mammalian host cells are well known and include, but are not limited to, those obtained from the genomes of viruses such as polyoma virus, fowlpox virus, adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, retroviruses, hepatitis-B virus and most preferably Simian Virus 40 (SV40). Other suitable mammalian promoters include heterologous mammalian promoters, for example, heat-shock promoters and the actin promoter.

Additional promoters which may be of interest in controlling LGR8 gene expression include, but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, *Nature* 290:304-10); the CMV promoter; the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto, *et al.*, 1980, *Cell* 22:787-97);

5 the herpes thymidine kinase promoter (Wagner *et al.*, 1981, *Proc. Natl. Acad. Sci. U.S.A.* 78:1444-45); the regulatory sequences of the metallothionine gene (Brinster *et al.*, 1982, *Nature* 296:39-42); prokaryotic expression vectors such as the beta-lactamase promoter (Villa-Kamaroff *et al.*, 1978, *Proc. Natl. Acad. Sci. U.S.A.*, 75:3727-31); or the tac promoter (DeBoer *et al.*, 1983, *Proc. Natl. Acad. Sci. U.S.A.*, 80:21-25). Also of interest

10 are the following animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: the elastase I gene control region which is active in pancreatic acinar cells (Swift *et al.*, 1984, *Cell* 38:639-46; Ornitz *et al.*, 1986, *Cold Spring Harbor Symp. Quant. Biol.* 50:399-409 (1986); MacDonald, 1987, *Hepatology* 7:425-515); the insulin gene control region which is active in pancreatic beta

15 cells (Hanahan, 1985, *Nature* 315:115-22); the immunoglobulin gene control region which is active in lymphoid cells (Grosschedl *et al.*, 1984, *Cell* 38:647-58; Adames *et al.*, 1985, *Nature* 318:533-38; Alexander *et al.*, 1987, *Mol. Cell. Biol.*, 7:1436-44); the mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells (Leder *et al.*, 1986, *Cell* 45:485-95); the albumin gene control region which is

20 active in liver (Pinkert *et al.*, 1987, *Genes and Devel.* 1:268-76); the alpha-feto-protein gene control region which is active in liver (Krumlauf *et al.*, 1985, *Mol. Cell. Biol.*, 5:1639-48; Hammer *et al.*, 1987, *Science* 235:53-58); the alpha 1-antitrypsin gene control region which is active in the liver (Kelsey *et al.*, 1987, *Genes and Devel.* 1:161-71); the beta-globin gene control region which is active in myeloid cells (Mogram *et al.*, 1985,

25 *Nature* 315:338-40; Kollias *et al.*, 1986, *Cell* 46:89-94); the myelin basic protein gene control region which is active in oligodendrocyte cells in the brain (Readhead *et al.*, 1987, *Cell* 48:703-12); the myosin light chain-2 gene control region which is active in skeletal muscle (Sani, 1985, *Nature* 314:283-86); and the gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason *et al.*, 1986, *Science*

30 234:1372-78).

An enhancer sequence may be inserted into the vector to increase the transcription of a DNA encoding an LGR8 polypeptide of the present invention by higher eukaryotes. Enhancers are cis-acting elements of DNA, usually about 10-300 bp in length, that act on the promoter to increase transcription. Enhancers are relatively orientation and position independent. They have been found 5' and 3' to the transcription unit. Several enhancer sequences available from mammalian genes are known (*e.g.*, globin, elastase, albumin, alpha-feto-protein and insulin). Typically, however, an enhancer from a virus will be used. The SV40 enhancer, the cytomegalovirus early promoter enhancer, the polyoma enhancer, and adenovirus enhancers are exemplary enhancing elements for the activation of eukaryotic promoters. While an enhancer may be spliced into the vector at a position 5' or 3' to an LGR8 nucleic acid molecule, it is typically located at a site 5' from the promoter.

Expression vectors of the invention may be constructed from a starting vector such as a commercially available vector. Such vectors may or may not contain all of the desired flanking sequences. Where one or more of the flanking sequences described herein are not already present in the vector, they may be individually obtained and ligated into the vector. Methods used for obtaining each of the flanking sequences are well known to one skilled in the art.

Preferred vectors for practicing this invention are those that are compatible with bacterial, insect, and mammalian host cells. Such vectors include, *inter alia*, pCRII, pCR3, and pcDNA3.1 (Invitrogen, San Diego, CA), pBSII (Stratagene, La Jolla, CA), pET15 (Novagen, Madison, WI), pGEX (Pharmacia Biotech, Piscataway, NJ), pEGFP-N2 (Clontech, Palo Alto, CA), pETL (BlueBacII, Invitrogen), pDSR-alpha (PCT Pub. No. WO 90/14363) and pFastBacDual (Gibco-BRL, Grand Island, NY).

Additional suitable vectors include, but are not limited to, cosmids, plasmids, or modified viruses, but it will be appreciated that the vector system must be compatible with the selected host cell. Such vectors include, but are not limited to plasmids such as Bluescript<sup>®</sup> plasmid derivatives (a high copy number ColE1-based phagemid; Stratagene Cloning Systems, La Jolla CA), PCR cloning plasmids designed for cloning Taq-amplified PCR products (*e.g.*, TOPO<sup>™</sup> TA Cloning<sup>®</sup> Kit and PCR2.1<sup>®</sup> plasmid



derivatives; Invitrogen), and mammalian, yeast or virus vectors such as a baculovirus expression system (pBacPAK plasmid derivatives; Clontech).

After the vector has been constructed and a nucleic acid molecule encoding an LGR8 polypeptide has been inserted into the proper site of the vector, the completed  
5 vector may be inserted into a suitable host cell for amplification and/or polypeptide expression. The transformation of an expression vector for an LGR8 polypeptide into a selected host cell may be accomplished by well known methods including methods such as transfection, infection, calcium chloride, electroporation, microinjection, lipofection, DEAE-dextran method, or other known techniques. The method selected will in part be a  
10 function of the type of host cell to be used. These methods and other suitable methods are well known to the skilled artisan, and are set forth, for example, in Sambrook *et al.*, *supra*.

Host cells may be prokaryotic host cells (such as *E. coli*) or eukaryotic host cells (such as a yeast, insect, or vertebrate cell). The host cell, when cultured under  
15 appropriate conditions, synthesizes an LGR8 polypeptide that can subsequently be collected from the culture medium (if the host cell secretes it into the medium) or directly from the host cell producing it (if it is not secreted). The selection of an appropriate host cell will depend upon various factors, such as desired expression levels, polypeptide modifications that are desirable or necessary for activity (such as glycosylation or  
20 phosphorylation) and ease of folding into a biologically active molecule.

A number of suitable host cells are known in the art and many are available from the American Type Culture Collection (ATCC), Manassas, VA. Examples include, but are not limited to, mammalian cells, such as Chinese hamster ovary cells (CHO), CHO DHFR(-) cells (Urlaub *et al.*, 1980, *Proc. Natl. Acad. Sci. U.S.A.* 97:4216-20), human  
25 embryonic kidney (HEK) 293 or 293T cells, or 3T3 cells. The selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening, product production, and purification are known in the art. Other suitable mammalian cell lines, are the monkey COS-1 and COS-7 cell lines, and the CV-1 cell line. Further exemplary mammalian host cells include primate cell lines and rodent cell lines,  
30 including transformed cell lines. Normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, as well as primary explants, are also suitable. Candidate cells

may be genotypically deficient in the selection gene, or may contain a dominantly acting selection gene. Other suitable mammalian cell lines include but are not limited to, mouse neuroblastoma N2A cells, HeLa, mouse L-929 cells, 3T3 lines derived from Swiss, Balb-c or NIH mice, BHK or HaK hamster cell lines. Each of these cell lines is known by and available to those skilled in the art of protein expression.

Similarly useful as host cells suitable for the present invention are bacterial cells. For example, the various strains of *E. coli* (e.g., HB101, DH5 $\alpha$ , DH10, and MC1061) are well-known as host cells in the field of biotechnology. Various strains of *B. subtilis*, *Pseudomonas spp.*, other *Bacillus spp.*, *Streptomyces spp.*, and the like may also be employed in this method.

Many strains of yeast cells known to those skilled in the art are also available as host cells for the expression of the polypeptides of the present invention. Preferred yeast cells include, for example, *Saccharomyces cerevisiae* and *Pichia pastoris*.

Additionally, where desired, insect cell systems may be utilized in the methods of the present invention. Such systems are described, for example, in Kitts *et al.*, 1993, *Biotechniques*, 14:810-17; Lucklow, 1993, *Curr. Opin. Biotechnol.* 4:564-72; and Lucklow *et al.*, 1993, *J. Virol.*, 67:4566-79. Preferred insect cells are Sf-9 and Hi5 (Invitrogen).

One may also use transgenic animals to express glycosylated LGR8 polypeptides. For example, one may use a transgenic milk-producing animal (a cow or goat, for example) and obtain the present glycosylated polypeptide in the animal milk. One may also use plants to produce LGR8 polypeptides, however, in general, the glycosylation occurring in plants is different from that produced in mammalian cells, and may result in a glycosylated product which is not suitable for human therapeutic use.

#### Polypeptide Production

Host cells comprising an LGR8 polypeptide expression vector may be cultured using standard media well known to the skilled artisan. The media will usually contain all nutrients necessary for the growth and survival of the cells. Suitable media for culturing *E. coli* cells include, for example, Luria Broth (LB) and/or Terrific Broth (TB). Suitable media for culturing eukaryotic cells include Roswell Park Memorial Institute

medium 1640 (RPMI 1640), Minimal Essential Medium (MEM) and/or Dulbecco's Modified Eagle Medium (DMEM), all of which may be supplemented with serum and/or growth factors as necessary for the particular cell line being cultured. A suitable medium for insect cultures is Grace's medium supplemented with yeastolate, lactalbumin hydrolysate, and/or fetal calf serum as necessary.

Typically, an antibiotic or other compound useful for selective growth of transfected or transformed cells is added as a supplement to the media. The compound to be used will be dictated by the selectable marker element present on the plasmid with which the host cell was transformed. For example, where the selectable marker element is kanamycin resistance, the compound added to the culture medium will be kanamycin. Other compounds for selective growth include ampicillin, tetracycline, and neomycin.

The amount of an LGR8 polypeptide produced by a host cell can be evaluated using standard methods known in the art. Such methods include, without limitation, Western blot analysis, SDS-polyacrylamide gel electrophoresis, non-denaturing gel electrophoresis, High Performance Liquid Chromatography (HPLC) separation, immunoprecipitation, and/or activity assays such as DNA binding gel shift assays.

If an LGR8 polypeptide has been designed to be secreted from the host cells, the majority of polypeptide may be found in the cell culture medium. If however, the LGR8 polypeptide is not secreted from the host cells, it will be present in the cytoplasm and/or the nucleus (for eukaryotic host cells) or in the cytosol (for gram-negative bacteria host cells).

For an LGR8 polypeptide situated in the host cell cytoplasm and/or nucleus (for eukaryotic host cells) or in the cytosol (for bacterial host cells), the intracellular material (including inclusion bodies for gram-negative bacteria) can be extracted from the host cell using any standard technique known to the skilled artisan. For example, the host cells can be lysed to release the contents of the periplasm/cytoplasm by French press, homogenization, and/or sonication followed by centrifugation.

If an LGR8 polypeptide has formed inclusion bodies in the cytosol, the inclusion bodies can often bind to the inner and/or outer cellular membranes and thus will be found primarily in the pellet material after centrifugation. The pellet material can then be treated at pH extremes or with a chaotropic agent such as a detergent, guanidine,

guanidine derivatives, urea, or urea derivatives in the presence of a reducing agent such as dithiothreitol at alkaline pH or tris carboxyethyl phosphine at acid pH to release, break apart, and solubilize the inclusion bodies. The solubilized LGR8 polypeptide can then be analyzed using gel electrophoresis, immunoprecipitation, or the like. If it is desired to isolate the LGR8 polypeptide, isolation may be accomplished using standard methods such as those described herein and in Marston *et al.*, 1990, *Meth. Enz.*, 182:264-75.

In some cases, an LGR8 polypeptide may not be biologically active upon isolation. Various methods for “refolding” or converting the polypeptide to its tertiary structure and generating disulfide linkages can be used to restore biological activity. Such methods include exposing the solubilized polypeptide to a pH usually above 7 and in the presence of a particular concentration of a chaotrope. The selection of chaotrope is very similar to the choices used for inclusion body solubilization, but usually the chaotrope is used at a lower concentration and is not necessarily the same as chaotropes used for the solubilization. In most cases the refolding/oxidation solution will also contain a reducing agent or the reducing agent plus its oxidized form in a specific ratio to generate a particular redox potential allowing for disulfide shuffling to occur in the formation of the protein’s cysteine bridges. Some of the commonly used redox couples include cysteine/cystamine, glutathione (GSH)/dithiobis GSH, cupric chloride, dithiothreitol(DTT)/dithiane DTT, and 2-2-mercaptoethanol(bME)/dithio-b(ME). In many instances, a cosolvent may be used or may be needed to increase the efficiency of the refolding, and the more common reagents used for this purpose include glycerol, polyethylene glycol of various molecular weights, arginine and the like.

If inclusion bodies are not formed to a significant degree upon expression of an LGR8 polypeptide, then the polypeptide will be found primarily in the supernatant after centrifugation of the cell homogenate. The polypeptide may be further isolated from the supernatant using methods such as those described herein.

The purification of an LGR8 polypeptide from solution can be accomplished using a variety of techniques. If the polypeptide has been synthesized such that it contains a tag such as Hexahistidine (LGR8 polypeptide/hexaHis) or other small peptide such as FLAG (Eastman Kodak Co., New Haven, CT) or *myc* (Invitrogen) at either its carboxyl- or amino-terminus, it may be purified in a one-step process by passing the

solution through an affinity column where the column matrix has a high affinity for the tag.

For example, polyhistidine binds with great affinity and specificity to nickel. Thus, an affinity column of nickel (such as the Qiagen® nickel columns) can be used for purification of LGR8 polypeptide/polyHis. See, e.g., *Current Protocols in Molecular Biology* § 10.11.8 (Ausubel *et al.*, eds., Green Publishers Inc. and Wiley and Sons 1993).

Additionally, LGR8 polypeptides may be purified through the use of a monoclonal antibody that is capable of specifically recognizing and binding to an LGR8 polypeptide.

Other suitable procedures for purification include, without limitation, affinity chromatography, immunoaffinity chromatography, ion exchange chromatography, molecular sieve chromatography, HPLC, electrophoresis (including native gel electrophoresis) followed by gel elution, and preparative isoelectric focusing ("Isoprime" machine/technique, Hoefer Scientific, San Francisco, CA). In some cases, two or more purification techniques may be combined to achieve increased purity.

LGR8 polypeptides may also be prepared by chemical synthesis methods (such as solid phase peptide synthesis) using techniques known in the art such as those set forth by Merrifield *et al.*, 1963, *J. Am. Chem. Soc.* 85:2149; Houghten *et al.*, 1985, *Proc Natl Acad. Sci. USA* 82:5132; and Stewart and Young, *Solid Phase Peptide Synthesis* (Pierce Chemical Co. 1984). Such polypeptides may be synthesized with or without a methionine on the amino-terminus. Chemically synthesized LGR8 polypeptides may be oxidized using methods set forth in these references to form disulfide bridges. Chemically synthesized LGR8 polypeptides are expected to have comparable biological activity to the corresponding LGR8 polypeptides produced recombinantly or purified from natural sources, and thus may be used interchangeably with a recombinant or natural LGR8 polypeptide.

Another means of obtaining LGR8 polypeptide is via purification from biological samples such as source tissues and/or fluids in which the LGR8 polypeptide is naturally found. Such purification can be conducted using methods for protein purification as described herein. The presence of the LGR8 polypeptide during purification may be

monitored, for example, using an antibody prepared against recombinantly produced LGR8 polypeptide or peptide fragments thereof.

5 A number of additional methods for producing nucleic acids and polypeptides are known in the art, and the methods can be used to produce polypeptides having specificity for LGR8 polypeptide. *See, e.g., Roberts et al., 1997, Proc. Natl. Acad. Sci. U.S.A.* 94:12297-303, which describes the production of fusion proteins between an mRNA and its encoded peptide. *See also, Roberts, 1999, Curr. Opin. Chem. Biol.* 3:268-73. Additionally, U.S. Patent No. 5,824,469 describes methods for obtaining oligonucleotides capable of carrying out a specific biological function. The procedure involves generating  
10 a heterogeneous pool of oligonucleotides, each having a 5' randomized sequence, a central preselected sequence, and a 3' randomized sequence. The resulting heterogeneous pool is introduced into a population of cells that do not exhibit the desired biological function. Subpopulations of the cells are then screened for those that exhibit a predetermined biological function. From that subpopulation, oligonucleotides capable of  
15 carrying out the desired biological function are isolated.

U.S. Patent Nos. 5,763,192; 5,814,476; 5,723,323; and 5,817,483 describe processes for producing peptides or polypeptides. This is done by producing stochastic genes or fragments thereof, and then introducing these genes into host cells which produce one or more proteins encoded by the stochastic genes. The host cells are then  
20 screened to identify those clones producing peptides or polypeptides having the desired activity.

Another method for producing peptides or polypeptides is described in PCT/US98/20094 (WO99/15650) filed by Athersys, Inc. Known as "Random Activation of Gene Expression for Gene Discovery" (RAGE-GD), the process involves the  
25 activation of endogenous gene expression or over-expression of a gene by *in situ* recombination methods. For example, expression of an endogenous gene is activated or increased by integrating a regulatory sequence into the target cell that is capable of activating expression of the gene by non-homologous or illegitimate recombination. The target DNA is first subjected to radiation, and a genetic promoter inserted. The promoter  
30 eventually locates a break at the front of a gene, initiating transcription of the gene. This results in expression of the desired peptide or polypeptide.

It will be appreciated that these methods can also be used to create comprehensive LGR8 polypeptide expression libraries, which can subsequently be used for high throughput phenotypic screening in a variety of assays, such as biochemical assays, cellular assays, and whole organism assays (*e.g.*, plant, mouse, etc.).

5

### Synthesis

It will be appreciated by those skilled in the art that the nucleic acid and polypeptide molecules described herein may be produced by recombinant and other means.

10

### Selective Binding Agents

The term “selective binding agent” refers to a molecule that has specificity for one or more LGR8 polypeptides. Suitable selective binding agents include, but are not limited to, antibodies and derivatives thereof, polypeptides, and small molecules. Suitable selective binding agents may be prepared using methods known in the art. An exemplary LGR8 polypeptide selective binding agent of the present invention is capable of binding a certain portion of the LGR8 polypeptide thereby inhibiting the binding of the polypeptide to an LGR8 polypeptide receptor.

15

Selective binding agents such as antibodies and antibody fragments that bind LGR8 polypeptides are within the scope of the present invention. The antibodies may be polyclonal including monospecific polyclonal; monoclonal (MAbs); recombinant; chimeric; humanized, such as complementarity-determining region (CDR)-grafted; human; single chain; and/or bispecific; as well as fragments; variants; or derivatives thereof. Antibody fragments include those portions of the antibody that bind to an epitope on the LGR8 polypeptide. Examples of such fragments include Fab and F(ab') fragments generated by enzymatic cleavage of full-length antibodies. Other binding fragments include those generated by recombinant DNA techniques, such as the expression of recombinant plasmids containing nucleic acid sequences encoding antibody variable regions.

20

25

Polyclonal antibodies directed toward an LGR8 polypeptide generally are produced in animals (*e.g.*, rabbits or mice) by means of multiple subcutaneous or

30

intraperitoneal injections of LGR8 polypeptide and an adjuvant. It may be useful to conjugate an LGR8 polypeptide to a carrier protein that is immunogenic in the species to be immunized, such as keyhole limpet hemocyanin, serum, albumin, bovine thyroglobulin, or soybean trypsin inhibitor. Also, aggregating agents such as alum are used to enhance the immune response. After immunization, the animals are bled and the serum is assayed for anti-LGR8 antibody titer.

Monoclonal antibodies directed toward LGR8 polypeptides are produced using any method that provides for the production of antibody molecules by continuous cell lines in culture. Examples of suitable methods for preparing monoclonal antibodies include the hybridoma methods of Kohler *et al.*, 1975, *Nature* 256:495-97 and the human B-cell hybridoma method (Kozbor, 1984, *J. Immunol.* 133:3001; Brodeur *et al.*, *Monoclonal Antibody Production Techniques and Applications* 51-63 (Marcel Dekker, Inc., 1987). Also provided by the invention are hybridoma cell lines that produce monoclonal antibodies reactive with LGR8 polypeptides.

Monoclonal antibodies of the invention may be modified for use as therapeutics. One embodiment is a "chimeric" antibody in which a portion of the heavy (H) and/or light (L) chain is identical with or homologous to a corresponding sequence in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is/are identical with or homologous to a corresponding sequence in antibodies derived from another species or belonging to another antibody class or subclass. Also included are fragments of such antibodies, so long as they exhibit the desired biological activity. See U.S. Patent No. 4,816,567; Morrison *et al.*, 1985, *Proc. Natl. Acad. Sci.* 81:6851-55.

In another embodiment, a monoclonal antibody of the invention is a "humanized" antibody. Methods for humanizing non-human antibodies are well known in the art. See U.S. Patent Nos. 5,585,089 and 5,693,762. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source that is non-human. Humanization can be performed, for example, using methods described in the art (Jones *et al.*, 1986, *Nature* 321:522-25; Riechmann *et al.*, 1998, *Nature* 332:323-27; Verhoeyen *et al.*, 1988, *Science* 239:1534-36), by substituting at least a portion of a rodent complementarity-determining region for the corresponding regions of a human antibody.



Also encompassed by the invention are human antibodies that bind LGR8 polypeptides. Using transgenic animals (*e.g.*, mice) that are capable of producing a repertoire of human antibodies in the absence of endogenous immunoglobulin production such antibodies are produced by immunization with an LGR8 polypeptide antigen (*i.e.*, having at least 6 contiguous amino acids), optionally conjugated to a carrier. *See, e.g.*, Jakobovits *et al.*, 1993, *Proc. Natl. Acad. Sci.* 90:2551-55; Jakobovits *et al.*, 1993, *Nature* 362:255-58; Bruggermann *et al.*, 1993, *Year in Immuno.* 7:33. In one method, such transgenic animals are produced by incapacitating the endogenous loci encoding the heavy and light immunoglobulin chains therein, and inserting loci encoding human heavy and light chain proteins into the genome thereof. Partially modified animals (*i.e.*, those having less than the full complement of modifications) are then cross-bred to obtain an animal having all of the desired immune system modifications. When administered an immunogen, these transgenic animals produce antibodies with human (rather than, *e.g.*, murine) amino acid sequences, including variable regions that are immunospecific for these antigens. *See* PCT App. Nos. PCT/US96/05928 and PCT/US93/06926. Additional methods are described in U.S. Patent No. 5,545,807, PCT App. Nos. PCT/US91/245 and PCT/GB89/01207, and in European Patent Nos. 546073B1 and 546073A1. Human antibodies can also be produced by the expression of recombinant DNA in host cells or by expression in hybridoma cells as described herein.

In an alternative embodiment, human antibodies can also be produced from phage-display libraries (Hoogenboom *et al.*, 1991, *J. Mol. Biol.* 227:381; Marks *et al.*, 1991, *J. Mol. Biol.* 222:581). These processes mimic immune selection through the display of antibody repertoires on the surface of filamentous bacteriophage, and subsequent selection of phage by their binding to an antigen of choice. One such technique is described in PCT App. No. PCT/US98/17364, which describes the isolation of high affinity and functional agonistic antibodies for MPL- and msk-receptors using such an approach.

Chimeric, CDR grafted, and humanized antibodies are typically produced by recombinant methods. Nucleic acids encoding the antibodies are introduced into host cells and expressed using materials and procedures described herein. In a preferred embodiment, the antibodies are produced in mammalian host cells, such as CHO cells.

Monoclonal (e.g., human) antibodies may be produced by the expression of recombinant DNA in host cells or by expression in hybridoma cells as described herein.

5 The anti-LGR8 antibodies of the invention may be employed in any known assay method, such as competitive binding assays, direct and indirect sandwich assays, and immunoprecipitation assays (Sola, *Monoclonal Antibodies: A Manual of Techniques* 147-158 (CRC Press, Inc., 1987)) for the detection and quantitation of LGR8 polypeptides. The antibodies will bind LGR8 polypeptides with an affinity that is appropriate for the assay method being employed.

10 For diagnostic applications, in certain embodiments, anti-LGR8 antibodies may be labeled with a detectable moiety. The detectable moiety can be any one that is capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as <sup>3</sup>H, <sup>14</sup>C, <sup>32</sup>P, <sup>35</sup>S, <sup>125</sup>I, <sup>99</sup>Tc, <sup>111</sup>In, or <sup>67</sup>Ga; a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin; or an enzyme, such as alkaline phosphatase,  $\beta$ -galactosidase, or  
15 horseradish peroxidase (Bayer, *et al.*, 1990, *Meth. Enz.* 184:138-63).

Competitive binding assays rely on the ability of a labeled standard (e.g., an LGR8 polypeptide, or an immunologically reactive portion thereof) to compete with the test sample analyte (an LGR8 polypeptide) for binding with a limited amount of anti-LGR8 antibody. The amount of an LGR8 polypeptide in the test sample is inversely  
20 proportional to the amount of standard that becomes bound to the antibodies. To facilitate determining the amount of standard that becomes bound, the antibodies typically are insolubilized before or after the competition, so that the standard and analyte that are bound to the antibodies may conveniently be separated from the standard and analyte that remain unbound.

25 Sandwich assays typically involve the use of two antibodies, each capable of binding to a different immunogenic portion, or epitope, of the protein to be detected and/or quantitated. In a sandwich assay, the test sample analyte is typically bound by a first antibody that is immobilized on a solid support, and thereafter a second antibody binds to the analyte, thus forming an insoluble three-part complex. See, e.g., U.S. Patent  
30 No. 4,376,110. The second antibody may itself be labeled with a detectable moiety (direct sandwich assays) or may be measured using an anti-immunoglobulin antibody

that is labeled with a detectable moiety (indirect sandwich assays). For example, one type of sandwich assay is an enzyme-linked immunosorbent assay (ELISA), in which case the detectable moiety is an enzyme.

The selective binding agents, including anti-LGR8 antibodies, are also useful for *in vivo* imaging. An antibody labeled with a detectable moiety may be administered to an animal, preferably into the bloodstream, and the presence and location of the labeled antibody in the host assayed. The antibody may be labeled with any moiety that is detectable in an animal, whether by nuclear magnetic resonance, radiology, or other detection means known in the art.

Selective binding agents of the invention, including antibodies, may be used as therapeutics. These therapeutic agents are generally agonists or antagonists, in that they either enhance or reduce, respectively, at least one of the biological activities of an LGR8 polypeptide. In one embodiment, antagonist antibodies of the invention are antibodies or binding fragments thereof which are capable of specifically binding to an LGR8 polypeptide and which are capable of inhibiting or eliminating the functional activity of an LGR8 polypeptide *in vivo* or *in vitro*. In preferred embodiments, the selective binding agent, *e.g.*, an antagonist antibody, will inhibit the functional activity of an LGR8 polypeptide by at least about 50%, and preferably by at least about 80%. In another embodiment, the selective binding agent may be an anti-LGR8 polypeptide antibody that is capable of interacting with an LGR8 polypeptide binding partner (a ligand or receptor) thereby inhibiting or eliminating LGR8 polypeptide activity *in vitro* or *in vivo*. Selective binding agents, including agonist and antagonist anti-LGR8 polypeptide antibodies, are identified by screening assays that are well known in the art.

The invention also relates to a kit comprising LGR8 selective binding agents (such as antibodies) and other reagents useful for detecting LGR8 polypeptide levels in biological samples. Such reagents may include a detectable label, blocking serum, positive and negative control samples, and detection reagents.

#### Microarrays

It will be appreciated that DNA microarray technology can be utilized in accordance with the present invention. DNA microarrays are miniature, high-density

arrays of nucleic acids positioned on a solid support, such as glass. Each cell or element within the array contains numerous copies of a single nucleic acid species that acts as a target for hybridization with a complementary nucleic acid sequence (*e.g.*, mRNA). In expression profiling using DNA microarray technology, mRNA is first extracted from a cell or tissue sample and then converted enzymatically to fluorescently labeled cDNA. This material is hybridized to the microarray and unbound cDNA is removed by washing. The expression of discrete genes represented on the array is then visualized by quantitating the amount of labeled cDNA that is specifically bound to each target nucleic acid molecule. In this way, the expression of thousands of genes can be quantitated in a high throughput, parallel manner from a single sample of biological material.

This high throughput expression profiling has a broad range of applications with respect to the LGR8 molecules of the invention, including, but not limited to: the identification and validation of LGR8 disease-related genes as targets for therapeutics; molecular toxicology of related LGR8 molecules and inhibitors thereof; stratification of populations and generation of surrogate markers for clinical trials; and enhancing related LGR8 polypeptide small molecule drug discovery by aiding in the identification of selective compounds in high throughput screens.

#### Chemical Derivatives

Chemically modified derivatives of LGR8 polypeptides may be prepared by one skilled in the art, given the disclosures described herein. LGR8 polypeptide derivatives are modified in a manner that is different – either in the type or location of the molecules naturally attached to the polypeptide. Derivatives may include molecules formed by the deletion of one or more naturally-attached chemical groups. The polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or other LGR8 polypeptide, may be modified by the covalent attachment of one or more polymers. For example, the polymer selected is typically water-soluble so that the protein to which it is attached does not precipitate in an aqueous environment, such as a physiological environment. Included within the scope of suitable polymers is a mixture

of polymers. Preferably, for therapeutic use of the end-product preparation, the polymer will be pharmaceutically acceptable.

The polymers each may be of any molecular weight and may be branched or unbranched. The polymers each typically have an average molecular weight of between about 2 kDa to about 100 kDa (the term "about" indicating that in preparations of a water-soluble polymer, some molecules will weigh more, some less, than the stated molecular weight). The average molecular weight of each polymer is preferably between about 5 kDa and about 50 kDa, more preferably between about 12 kDa and about 40 kDa and most preferably between about 20 kDa and about 35 kDa.

Suitable water-soluble polymers or mixtures thereof include, but are not limited to, N-linked or O-linked carbohydrates, sugars, phosphates, polyethylene glycol (PEG) (including the forms of PEG that have been used to derivatize proteins, including mono- (C<sub>1</sub>-C<sub>10</sub>), alkoxy-, or aryloxy-polyethylene glycol), monomethoxy-polyethylene glycol, dextran (such as low molecular weight dextran of, for example, about 6 kD), cellulose, or other carbohydrate based polymers, poly-(N-vinyl pyrrolidone) polyethylene glycol, propylene glycol homopolymers, polypropylene oxide/ethylene oxide co-polymers, polyoxyethylated polyols (e.g., glycerol), and polyvinyl alcohol. Also encompassed by the present invention are bifunctional crosslinking molecules that may be used to prepare covalently attached LGR8 polypeptide multimers.

In general, chemical derivatization may be performed under any suitable condition used to react a protein with an activated polymer molecule. Methods for preparing chemical derivatives of polypeptides will generally comprise the steps of: (a) reacting the polypeptide with the activated polymer molecule (such as a reactive ester or aldehyde derivative of the polymer molecule) under conditions whereby the polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or other LGR8 polypeptide, becomes attached to one or more polymer molecules, and (b) obtaining the reaction products. The optimal reaction conditions will be determined based on known parameters and the desired result. For example, the larger the ratio of polymer molecules to protein, the greater the percentage

of attached polymer molecule. In one embodiment, the LGR8 polypeptide derivative may have a single polymer molecule moiety at the amino-terminus. *See, e.g.*, U.S. Patent No. 5,234,784.

5 The pegylation of a polypeptide may be specifically carried out using any of the pegylation reactions known in the art. Such reactions are described, for example, in the following references: Francis *et al.*, 1992, *Focus on Growth Factors* 3:4-10; European Patent Nos. 0154316 and 0401384; and U.S. Patent No. 4,179,337. For example, pegylation may be carried out via an acylation reaction or an alkylation reaction with a reactive polyethylene glycol molecule (or an analogous reactive water-soluble polymer) as described herein. For the acylation reactions, a selected polymer should have a single reactive ester group. For reductive alkylation, a selected polymer should have a single reactive aldehyde group. A reactive aldehyde is, for example, polyethylene glycol propionaldehyde, which is water stable, or mono C<sub>1</sub>-C<sub>10</sub> alkoxy or aryloxy derivatives thereof (*see* U.S. Patent No. 5,252,714).

15 In another embodiment, LGR8 polypeptides may be chemically coupled to biotin. The biotin/LGR8 polypeptide molecules are then allowed to bind to avidin, resulting in tetravalent avidin/biotin/LGR8 polypeptide molecules. LGR8 polypeptides may also be covalently coupled to dinitrophenol (DNP) or trinitrophenol (TNP) and the resulting conjugates precipitated with anti-DNP or anti-TNP-IgM to form decameric conjugates with a valency of 10.

20 Generally, conditions that may be alleviated or modulated by the administration of the present LGR8 polypeptide derivatives include those described herein for LGR8 polypeptides. However, the LGR8 polypeptide derivatives disclosed herein may have additional activities, enhanced or reduced biological activity, or other characteristics, such as increased or decreased half-life, as compared to the non-derivatized molecules.

#### Genetically Engineered Non-Human Animals

25 Additionally included within the scope of the present invention are non-human animals such as mice, rats, or other rodents; rabbits, goats, sheep, or other farm animals, in which the genes encoding native LGR8 polypeptide have been disrupted (*i.e.*, “knocked out”) such that the level of expression of LGR8 polypeptide is significantly

decreased or completely abolished. Such animals may be prepared using techniques and methods such as those described in U.S. Patent No. 5,557,032.

The present invention further includes non-human animals such as mice, rats, or other rodents; rabbits, goats, sheep, or other farm animals, in which either the native form of an LGR8 gene for that animal or a heterologous LGR8 gene is over-expressed by the animal, thereby creating a “transgenic” animal. Such transgenic animals may be prepared using well known methods such as those described in U.S. Patent No 5,489,743 and PCT Pub. No. WO 94/28122.

The present invention further includes non-human animals in which the promoter for one or more of the LGR8 polypeptides of the present invention is either activated or inactivated (*e.g.*, by using homologous recombination methods) to alter the level of expression of one or more of the native LGR8 polypeptides.

These non-human animals may be used for drug candidate screening. In such screening, the impact of a drug candidate on the animal may be measured. For example, drug candidates may decrease or increase the expression of the LGR8 gene. In certain embodiments, the amount of LGR8 polypeptide that is produced may be measured after the exposure of the animal to the drug candidate. Additionally, in certain embodiments, one may detect the actual impact of the drug candidate on the animal. For example, over-expression of a particular gene may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate’s ability to decrease expression of the gene or its ability to prevent or inhibit a pathological condition. In other examples, the production of a particular metabolic product such as a fragment of a polypeptide, may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate’s ability to decrease the production of such a metabolic product or its ability to prevent or inhibit a pathological condition.

#### Assaying for Other Modulators of LGR8 Polypeptide Activity

In some situations, it may be desirable to identify molecules that are modulators, *i.e.*, agonists or antagonists, of the activity of LGR8 polypeptide. Natural or synthetic molecules that modulate LGR8 polypeptide may be identified using one or more screening assays, such as those described herein. Such molecules may be administered

either in an *ex vivo* manner or in an *in vivo* manner by injection, or by oral delivery, implantation device, or the like.

“Test molecule” refers to a molecule that is under evaluation for the ability to modulate (*i.e.*, increase or decrease) the activity of an LGR8 polypeptide. Most commonly, a test molecule will interact directly with an LGR8 polypeptide. However, it is also contemplated that a test molecule may also modulate LGR8 polypeptide activity indirectly, such as by affecting LGR8 gene expression, or by binding to an LGR8 polypeptide binding partner (*e.g.*, receptor or ligand). In one embodiment, a test molecule will bind to an LGR8 polypeptide with an affinity constant of at least about  $10^{-6}$  M, preferably about  $10^{-8}$  M, more preferably about  $10^{-9}$  M, and even more preferably about  $10^{-10}$  M.

Methods for identifying compounds that interact with LGR8 polypeptides are encompassed by the present invention. In certain embodiments, an LGR8 polypeptide is incubated with a test molecule under conditions that permit the interaction of the test molecule with an LGR8 polypeptide, and the extent of the interaction is measured. The test molecule can be screened in a substantially purified form or in a crude mixture.

In certain embodiments, an LGR8 polypeptide agonist or antagonist may be a protein, peptide, carbohydrate, lipid, or small molecular weight molecule that interacts with LGR8 polypeptide to regulate its activity. Molecules which regulate LGR8 polypeptide expression include nucleic acids which are complementary to nucleic acids encoding an LGR8 polypeptide, or are complementary to nucleic acids sequences which direct or control the expression of LGR8 polypeptide, and which act as anti-sense regulators of expression.

Once a test molecule has been identified as interacting with an LGR8 polypeptide, the molecule may be further evaluated for its ability to increase or decrease LGR8 polypeptide activity. The measurement of the interaction of a test molecule with LGR8 polypeptide may be carried out in several formats, including cell-based binding assays, membrane binding assays, solution-phase assays, and immunoassays. In general, a test molecule is incubated with an LGR8 polypeptide for a specified period of time, and LGR8 polypeptide activity is determined by one or more assays for measuring biological activity.



The interaction of test molecules with LGR8 polypeptides may also be assayed directly using polyclonal or monoclonal antibodies in an immunoassay. Alternatively, modified forms of LGR8 polypeptides containing epitope tags as described herein may be used in solution and immunoassays.

5 In the event that LGR8 polypeptides display biological activity through an interaction with a binding partner (*e.g.*, a receptor or a ligand), a variety of *in vitro* assays may be used to measure the binding of an LGR8 polypeptide to the corresponding binding partner (such as a selective binding agent, receptor, or ligand). These assays may be used to screen test molecules for their ability to increase or decrease the rate and/or the  
10 extent of binding of an LGR8 polypeptide to its binding partner. In one assay, an LGR8 polypeptide is immobilized in the wells of a microtiter plate. Radiolabeled LGR8 polypeptide binding partner (for example, iodinated LGR8 polypeptide binding partner) and a test molecule can then be added either one at a time (in either order) or simultaneously to the wells. After incubation, the wells can be washed and counted for  
15 radioactivity, using a scintillation counter, to determine the extent to which the binding partner bound to the LGR8 polypeptide. Typically, a molecule will be tested over a range of concentrations, and a series of control wells lacking one or more elements of the test assays can be used for accuracy in the evaluation of the results. An alternative to this method involves reversing the “positions” of the proteins, *i.e.*, immobilizing LGR8  
20 polypeptide binding partner to the microtiter plate wells, incubating with the test molecule and radiolabeled LGR8 polypeptide, and determining the extent of LGR8 polypeptide binding. *See, e.g., Current Protocols in Molecular Biology*, chap. 18 (Ausubel *et al.*, eds., Green Publishers Inc. and Wiley and Sons 1995).

As an alternative to radiolabeling, an LGR8 polypeptide or its binding partner  
25 may be conjugated to biotin, and the presence of biotinylated protein can then be detected using streptavidin linked to an enzyme, such as horse radish peroxidase (HRP) or alkaline phosphatase (AP), which can be detected colorometrically, or by fluorescent tagging of streptavidin. An antibody directed to an LGR8 polypeptide or to an LGR8 polypeptide binding partner, and which is conjugated to biotin, may also be used for purposes of  
30 detection following incubation of the complex with enzyme-linked streptavidin linked to AP or HRP.

An LGR8 polypeptide or an LGR8 polypeptide binding partner can also be immobilized by attachment to agarose beads, acrylic beads, or other types of such inert solid phase substrates. The substrate-protein complex can be placed in a solution containing the complementary protein and the test compound. After incubation, the beads can be precipitated by centrifugation, and the amount of binding between an LGR8 polypeptide and its binding partner can be assessed using the methods described herein. Alternatively, the substrate-protein complex can be immobilized in a column with the test molecule and complementary protein passing through the column. The formation of a complex between an LGR8 polypeptide and its binding partner can then be assessed using any of the techniques described herein (e.g., radiolabelling or antibody binding).

Another *in vitro* assay that is useful for identifying a test molecule that increases or decreases the formation of a complex between an LGR8 polypeptide binding protein and an LGR8 polypeptide binding partner is a surface plasmon resonance detector system such as the BIAcore assay system (Pharmacia, Piscataway, NJ). The BIAcore system is utilized as specified by the manufacturer. This assay essentially involves the covalent binding of either LGR8 polypeptide or an LGR8 polypeptide binding partner to a dextran-coated sensor chip that is located in a detector. The test compound and the other complementary protein can then be injected, either simultaneously or sequentially, into the chamber containing the sensor chip. The amount of complementary protein that binds can be assessed based on the change in molecular mass that is physically associated with the dextran-coated side of the sensor chip, with the change in molecular mass being measured by the detector system.

In some cases, it may be desirable to evaluate two or more test compounds together for their ability to increase or decrease the formation of a complex between an LGR8 polypeptide and an LGR8 polypeptide binding partner. In these cases, the assays set forth herein can be readily modified by adding such additional test compound(s) either simultaneously with, or subsequent to, the first test compound. The remainder of the steps in the assay are as set forth herein.

*In vitro* assays such as those described herein may be used advantageously to screen large numbers of compounds for an effect on the formation of a complex between an LGR8 polypeptide and LGR8 polypeptide binding partner. The assays may be

automated to screen compounds generated in phage display, synthetic peptide, and chemical synthesis libraries.

Compounds which increase or decrease the formation of a complex between an LGR8 polypeptide and an LGR8 polypeptide binding partner may also be screened in cell culture using cells and cell lines expressing either LGR8 polypeptide or LGR8 polypeptide binding partner. Cells and cell lines may be obtained from any mammal, but preferably will be from human or other primate, canine, or rodent sources. The binding of an LGR8 polypeptide to cells expressing LGR8 polypeptide binding partner at the surface is evaluated in the presence or absence of test molecules, and the extent of binding may be determined by, for example, flow cytometry using a biotinylated antibody to an LGR8 polypeptide binding partner. Cell culture assays can be used advantageously to further evaluate compounds that score positive in protein binding assays described herein.

Cell cultures can also be used to screen the impact of a drug candidate. For example, drug candidates may decrease or increase the expression of the LGR8 gene. In certain embodiments, the amount of LGR8 polypeptide or an LGR8 polypeptide fragment that is produced may be measured after exposure of the cell culture to the drug candidate. In certain embodiments, one may detect the actual impact of the drug candidate on the cell culture. For example, the over-expression of a particular gene may have a particular impact on the cell culture. In such cases, one may test a drug candidate's ability to increase or decrease the expression of the gene or its ability to prevent or inhibit a particular impact on the cell culture. In other examples, the production of a particular metabolic product such as a fragment of a polypeptide, may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate's ability to decrease the production of such a metabolic product in a cell culture.

A yeast two-hybrid system (Chien *et al.*, 1991, *Proc. Natl. Acad. Sci. U.S.A.* 88:9578-83) can be used to identify novel polypeptides that bind to, or interact with, LGR8 polypeptides. As an example, hybrid constructs comprising DNA encoding a cytoplasmic domain of an LGR8 polypeptide fused to a yeast GAL4-DNA binding domain may be used as a two-hybrid bait plasmid. Positive clones emerging from the screening may be characterized further to identify interacting proteins.

### Internalizing Proteins

The *tat* protein sequence (from HIV) can be used to internalize proteins into a cell. See, e.g., Falwell *et al.*, 1994, *Proc. Natl. Acad. Sci. U.S.A.* 91:664-68. For example, an 11 amino acid sequence (Y-G-R-K-K-R-R-Q-R-R-R; SEQ ID NO: 25) of the HIV *tat* protein (termed the “protein transduction domain,” or TAT PDT) has been described as mediating delivery across the cytoplasmic membrane and the nuclear membrane of a cell. See Schwarze *et al.*, 1999, *Science* 285:1569-72; and Nagahara *et al.*, 1998, *Nat. Med.* 4:1449-52. In these procedures, FITC-constructs (FITC-labeled G-G-G-Y-G-R-K-K-R-R-Q-R-R-R; SEQ ID NO: 26), which penetrate tissues following intraperitoneal administration, are prepared, and the binding of such constructs to cells is detected by fluorescence-activated cell sorting (FACS) analysis. Cells treated with a *tat*- $\beta$ -gal fusion protein will demonstrate  $\beta$ -gal activity. Following injection, expression of such a construct can be detected in a number of tissues, including liver, kidney, lung, heart, and brain tissue. It is believed that such constructs undergo some degree of unfolding in order to enter the cell, and as such, may require a refolding following entry into the cell.

It will thus be appreciated that the *tat* protein sequence may be used to internalize a desired polypeptide into a cell. For example, using the *tat* protein sequence, an LGR8 antagonist (such as an anti-LGR8 selective binding agent, small molecule, soluble receptor, or antisense oligonucleotide) can be administered intracellularly to inhibit the activity of an LGR8 molecule. As used herein, the term “LGR8 molecule” refers to both LGR8 nucleic acid molecules and LGR8 polypeptides as defined herein. Where desired, the LGR8 protein itself may also be internally administered to a cell using these procedures. See also, Straus, 1999, *Science* 285:1466-67.

### Cell Source Identification Using LGR8 Polypeptide

In accordance with certain embodiments of the invention, it may be useful to be able to determine the source of a certain cell type associated with an LGR8 polypeptide. For example, it may be useful to determine the origin of a disease or pathological condition as an aid in selecting an appropriate therapy. In certain embodiments, nucleic

acids encoding an LGR8 polypeptide can be used as a probe to identify cells described herein by screening the nucleic acids of the cells with such a probe. In other embodiments, one may use anti-LGR8 polypeptide antibodies to test for the presence of LGR8 polypeptide in cells, and thus, determine if such cells are of the types described herein.

#### LGR8 Polypeptide Compositions and Administration

Therapeutic compositions are within the scope of the present invention. Such LGR8 polypeptide pharmaceutical compositions may comprise a therapeutically effective amount of an LGR8 polypeptide or an LGR8 nucleic acid molecule in admixture with a pharmaceutically or physiologically acceptable formulation agent selected for suitability with the mode of administration. Pharmaceutical compositions may comprise a therapeutically effective amount of one or more LGR8 polypeptide selective binding agents in admixture with a pharmaceutically or physiologically acceptable formulation agent selected for suitability with the mode of administration.

Acceptable formulation materials preferably are nontoxic to recipients at the dosages and concentrations employed.

The pharmaceutical composition may contain formulation materials for modifying, maintaining, or preserving, for example, the pH, osmolarity, viscosity, clarity, color, isotonicity, odor, sterility, stability, rate of dissolution or release, adsorption, or penetration of the composition. Suitable formulation materials include, but are not limited to, amino acids (such as glycine, glutamine, asparagine, arginine, or lysine), antimicrobials, antioxidants (such as ascorbic acid, sodium sulfite, or sodium hydrogen-sulfite), buffers (such as borate, bicarbonate, Tris-HCl, citrates, phosphates, or other organic acids), bulking agents (such as mannitol or glycine), chelating agents (such as ethylenediamine tetraacetic acid (EDTA)), complexing agents (such as caffeine, polyvinylpyrrolidone, beta-cyclodextrin, or hydroxypropyl-beta-cyclodextrin), fillers, monosaccharides, disaccharides, and other carbohydrates (such as glucose, mannose, or dextrans), proteins (such as serum albumin, gelatin, or immunoglobulins), coloring, flavoring and diluting agents, emulsifying agents, hydrophilic polymers (such as polyvinylpyrrolidone), low molecular weight polypeptides, salt-forming counterions

(such as sodium), preservatives (such as benzalkonium chloride, benzoic acid, salicylic acid, thimerosal, phenethyl alcohol, methylparaben, propylparaben, chlorhexidine, sorbic acid, or hydrogen peroxide), solvents (such as glycerin, propylene glycol, or polyethylene glycol), sugar alcohols (such as mannitol or sorbitol), suspending agents, surfactants or wetting agents (such as pluronics; PEG; sorbitan esters; polysorbates such as polysorbate 20 or polysorbate 80; triton; tromethamine; lecithin; cholesterol or tyloxapal), stability enhancing agents (such as sucrose or sorbitol), tonicity enhancing agents (such as alkali metal halides – preferably sodium or potassium chloride – or mannitol sorbitol), delivery vehicles, diluents, excipients and/or pharmaceutical adjuvants. *See Remington's Pharmaceutical Sciences* (18th Ed., A.R. Gennaro, ed., Mack Publishing Company 1990).

The optimal pharmaceutical composition will be determined by a skilled artisan depending upon, for example, the intended route of administration, delivery format, and desired dosage. *See, e.g., Remington's Pharmaceutical Sciences, supra.* Such compositions may influence the physical state, stability, rate of *in vivo* release, and rate of *in vivo* clearance of the LGR8 molecule.

The primary vehicle or carrier in a pharmaceutical composition may be either aqueous or non-aqueous in nature. For example, a suitable vehicle or carrier for injection may be water, physiological saline solution, or artificial cerebrospinal fluid, possibly supplemented with other materials common in compositions for parenteral administration. Neutral buffered saline or saline mixed with serum albumin are further exemplary vehicles. Other exemplary pharmaceutical compositions comprise Tris buffer of about pH 7.0-8.5, or acetate buffer of about pH 4.0-5.5, which may further include sorbitol or a suitable substitute. In one embodiment of the present invention, LGR8 polypeptide compositions may be prepared for storage by mixing the selected composition having the desired degree of purity with optional formulation agents (*Remington's Pharmaceutical Sciences, supra*) in the form of a lyophilized cake or an aqueous solution. Further, the LGR8 polypeptide product may be formulated as a lyophilizate using appropriate excipients such as sucrose.

The LGR8 polypeptide pharmaceutical compositions can be selected for parenteral delivery. Alternatively, the compositions may be selected for inhalation or for delivery through the digestive tract, such as orally. The preparation of such

pharmaceutically acceptable compositions is within the skill of the art.

The formulation components are present in concentrations that are acceptable to the site of administration. For example, buffers are used to maintain the composition at physiological pH or at a slightly lower pH, typically within a pH range of from about 5 to about 8.

When parenteral administration is contemplated, the therapeutic compositions for use in this invention may be in the form of a pyrogen-free, parenterally acceptable, aqueous solution comprising the desired LGR8 molecule in a pharmaceutically acceptable vehicle. A particularly suitable vehicle for parenteral injection is sterile distilled water in which an LGR8 molecule is formulated as a sterile, isotonic solution, properly preserved. Yet another preparation can involve the formulation of the desired molecule with an agent, such as injectable microspheres, bio-erodible particles, polymeric compounds (such as polylactic acid or polyglycolic acid), beads, or liposomes, that provides for the controlled or sustained release of the product which may then be delivered via a depot injection. Hyaluronic acid may also be used, and this may have the effect of promoting sustained duration in the circulation. Other suitable means for the introduction of the desired molecule include implantable drug delivery devices.

In one embodiment, a pharmaceutical composition may be formulated for inhalation. For example, LGR8 polypeptide may be formulated as a dry powder for inhalation. LGR8 polypeptide or nucleic acid molecule inhalation solutions may also be formulated with a propellant for aerosol delivery. In yet another embodiment, solutions may be nebulized. Pulmonary administration is further described in PCT Pub. No. WO 94/20069, which describes the pulmonary delivery of chemically modified proteins.

It is also contemplated that certain formulations may be administered orally. In one embodiment of the present invention, LGR8 polypeptides that are administered in this fashion can be formulated with or without those carriers customarily used in the compounding of solid dosage forms such as tablets and capsules. For example, a capsule may be designed to release the active portion of the formulation at the point in the gastrointestinal tract when bioavailability is maximized and pre-systemic degradation is minimized. Additional agents can be included to facilitate absorption of the LGR8

polypeptide. Diluents, flavorings, low melting point waxes, vegetable oils, lubricants, suspending agents, tablet disintegrating agents, and binders may also be employed.

Another pharmaceutical composition may involve an effective quantity of LGR8 polypeptides in a mixture with non-toxic excipients that are suitable for the manufacture of tablets. By dissolving the tablets in sterile water, or another appropriate vehicle, solutions can be prepared in unit-dose form. Suitable excipients include, but are not limited to, inert diluents, such as calcium carbonate, sodium carbonate or bicarbonate, lactose, or calcium phosphate; or binding agents, such as starch, gelatin, or acacia; or lubricating agents such as magnesium stearate, stearic acid, or talc.

Additional LGR8 polypeptide pharmaceutical compositions will be evident to those skilled in the art, including formulations involving LGR8 polypeptides in sustained- or controlled-delivery formulations. Techniques for formulating a variety of other sustained- or controlled-delivery means, such as liposome carriers, bio-erodible microparticles or porous beads and depot injections, are also known to those skilled in the art. *See, e.g.*, PCT/US93/00829, which describes the controlled release of porous polymeric microparticles for the delivery of pharmaceutical compositions.

Additional examples of sustained-release preparations include semipermeable polymer matrices in the form of shaped articles, *e.g.* films, or microcapsules. Sustained release matrices may include polyesters, hydrogels, polylactides (U.S. Patent No. 3,773,919 and European Patent No. 058481), copolymers of L-glutamic acid and gamma ethyl-L-glutamate (Sidman *et al.*, 1983, *Biopolymers* 22:547-56), poly(2-hydroxyethyl-methacrylate) (Langer *et al.*, 1981, *J. Biomed. Mater. Res.* 15:167-277 and Langer, 1982, *Chem. Tech.* 12:98-105), ethylene vinyl acetate (Langer *et al.*, *supra*) or poly-D(-)-3-hydroxybutyric acid (European Patent No. 133988). Sustained-release compositions may also include liposomes, which can be prepared by any of several methods known in the art. *See, e.g.*, Eppstein *et al.*, 1985, *Proc. Natl. Acad. Sci. USA* 82:3688-92; and European Patent Nos. 036676, 088046, and 143949.

The LGR8 pharmaceutical composition to be used for *in vivo* administration typically must be sterile. This may be accomplished by filtration through sterile filtration membranes. Where the composition is lyophilized, sterilization using this method may be conducted either prior to, or following, lyophilization and reconstitution. The



composition for parenteral administration may be stored in lyophilized form or in a solution. In addition, parenteral compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

5           Once the pharmaceutical composition has been formulated, it may be stored in sterile vials as a solution, suspension, gel, emulsion, solid, or as a dehydrated or lyophilized powder. Such formulations may be stored either in a ready-to-use form or in a form (*e.g.*, lyophilized) requiring reconstitution prior to administration.

10           In a specific embodiment, the present invention is directed to kits for producing a single-dose administration unit. The kits may each contain both a first container having a dried protein and a second container having an aqueous formulation. Also included within the scope of this invention are kits containing single and multi-chambered pre-filled syringes (*e.g.*, liquid syringes and lyosyringes).

15           The effective amount of an LGR8 pharmaceutical composition to be employed therapeutically will depend, for example, upon the therapeutic context and objectives. One skilled in the art will appreciate that the appropriate dosage levels for treatment will thus vary depending, in part, upon the molecule delivered, the indication for which the LGR8 molecule is being used, the route of administration, and the size (body weight, body surface, or organ size) and condition (the age and general health) of the patient.  
20           Accordingly, the clinician may titer the dosage and modify the route of administration to obtain the optimal therapeutic effect. A typical dosage may range from about 0.1  $\mu\text{g/kg}$  to up to about 100 mg/kg or more, depending on the factors mentioned above. In other embodiments, the dosage may range from 0.1  $\mu\text{g/kg}$  up to about 100 mg/kg; or 1  $\mu\text{g/kg}$  up to about 100 mg/kg; or 5  $\mu\text{g/kg}$  up to about 100 mg/kg.

25           The frequency of dosing will depend upon the pharmacokinetic parameters of the LGR8 molecule in the formulation being used. Typically, a clinician will administer the composition until a dosage is reached that achieves the desired effect. The composition may therefore be administered as a single dose, as two or more doses (which may or may not contain the same amount of the desired molecule) over time, or as a continuous  
30           infusion via an implantation device or catheter. Further refinement of the appropriate dosage is routinely made by those of ordinary skill in the art and is within the ambit of

tasks routinely performed by them. Appropriate dosages may be ascertained through use of appropriate dose-response data.

5 The route of administration of the pharmaceutical composition is in accord with known methods, *e.g.*, orally; through injection by intravenous, intraperitoneal, intracerebral (intraparenchymal), intracerebroventricular, intramuscular, intraocular, intraarterial, intraportal, or intralesional routes; by sustained release systems; or by implantation devices. Where desired, the compositions may be administered by bolus injection or continuously by infusion, or by implantation device.

10 Alternatively or additionally, the composition may be administered locally via implantation of a membrane, sponge, or other appropriate material onto which the desired molecule has been absorbed or encapsulated. Where an implantation device is used, the device may be implanted into any suitable tissue or organ, and delivery of the desired molecule may be via diffusion, timed-release bolus, or continuous administration.

15 In some cases, it may be desirable to use LGR8 polypeptide pharmaceutical compositions in an *ex vivo* manner. In such instances, cells, tissues, or organs that have been removed from the patient are exposed to LGR8 polypeptide pharmaceutical compositions after which the cells, tissues, or organs are subsequently implanted back into the patient.

20 In other cases, an LGR8 polypeptide can be delivered by implanting certain cells that have been genetically engineered, using methods such as those described herein, to express and secrete the LGR8 polypeptide. Such cells may be animal or human cells, and may be autologous, heterologous, or xenogeneic. Optionally, the cells may be immortalized. In order to decrease the chance of an immunological response, the cells may be encapsulated to avoid infiltration of surrounding tissues. The encapsulation  
25 materials are typically biocompatible, semi-permeable polymeric enclosures or membranes that allow the release of the protein product(s) but prevent the destruction of the cells by the patient's immune system or by other detrimental factors from the surrounding tissues.

30 Additional embodiments of the present invention relate to cells and methods (*e.g.*, homologous recombination and/or other recombinant production methods) for both the *in vitro* production of therapeutic polypeptides and for the production and delivery of

therapeutic polypeptides by gene therapy or cell therapy. Homologous and other recombination methods may be used to modify a cell that contains a normally transcriptionally-silent LGR8 gene, or an under-expressed gene, and thereby produce a cell which expresses therapeutically efficacious amounts of LGR8 polypeptides.

5 Homologous recombination is a technique originally developed for targeting genes to induce or correct mutations in transcriptionally active genes. Kucherlapati, 1989, *Prog. in Nucl. Acid Res. & Mol. Biol.* 36:301. The basic technique was developed as a method for introducing specific mutations into specific regions of the mammalian genome (Thomas *et al.*, 1986, *Cell* 44:419-28; Thomas and Capecchi, 1987, *Cell* 51:503-10 12; Doetschman *et al.*, 1988, *Proc. Natl. Acad. Sci. U.S.A.* 85:8583-87) or to correct specific mutations within defective genes (Doetschman *et al.*, 1987, *Nature* 330:576-78). Exemplary homologous recombination techniques are described in U.S. Patent No. 5,272,071; European Patent Nos. 9193051 and 505500; PCT/US90/07642, and PCT Pub No. WO 91/09955).

15 Through homologous recombination, the DNA sequence to be inserted into the genome can be directed to a specific region of the gene of interest by attaching it to targeting DNA. The targeting DNA is a nucleotide sequence that is complementary (homologous) to a region of the genomic DNA. Small pieces of targeting DNA that are complementary to a specific region of the genome are put in contact with the parental strand during the DNA replication process. It is a general property of DNA that has been 20 inserted into a cell to hybridize, and therefore, recombine with other pieces of endogenous DNA through shared homologous regions. If this complementary strand is attached to an oligonucleotide that contains a mutation or a different sequence or an additional nucleotide, it too is incorporated into the newly synthesized strand as a result 25 of the recombination. As a result of the proofreading function, it is possible for the new sequence of DNA to serve as the template. Thus, the transferred DNA is incorporated into the genome.

Attached to these pieces of targeting DNA are regions of DNA that may interact with or control the expression of an LGR8 polypeptide, *e.g.*, flanking sequences. For 30 example, a promoter/enhancer element, a suppressor, or an exogenous transcription modulatory element is inserted in the genome of the intended host cell in proximity and

orientation sufficient to influence the transcription of DNA encoding the desired LGR8 polypeptide. The control element controls a portion of the DNA present in the host cell genome. Thus, the expression of the desired LGR8 polypeptide may be achieved not by transfection of DNA that encodes the LGR8 gene itself, but rather by the use of targeting DNA (containing regions of homology with the endogenous gene of interest) coupled with DNA regulatory segments that provide the endogenous gene sequence with recognizable signals for transcription of an LGR8 gene.

In an exemplary method, the expression of a desired targeted gene in a cell (*i.e.*, a desired endogenous cellular gene) is altered via homologous recombination into the cellular genome at a preselected site, by the introduction of DNA that includes at least a regulatory sequence, an exon, and a splice donor site. These components are introduced into the chromosomal (genomic) DNA in such a manner that this, in effect, results in the production of a new transcription unit (in which the regulatory sequence, the exon, and the splice donor site present in the DNA construct are operatively linked to the endogenous gene). As a result of the introduction of these components into the chromosomal DNA, the expression of the desired endogenous gene is altered.

Altered gene expression, as described herein, encompasses activating (or causing to be expressed) a gene which is normally silent (unexpressed) in the cell as obtained, as well as increasing the expression of a gene which is not expressed at physiologically significant levels in the cell as obtained. The embodiments further encompass changing the pattern of regulation or induction such that it is different from the pattern of regulation or induction that occurs in the cell as obtained, and reducing (including eliminating) the expression of a gene which is expressed in the cell as obtained.

One method by which homologous recombination can be used to increase, or cause, LGR8 polypeptide production from a cell's endogenous LGR8 gene involves first using homologous recombination to place a recombination sequence from a site-specific recombination system (*e.g.*, Cre/loxP, FLP/FRT) (Sauer, 1994, *Curr. Opin. Biotechnol.*, 5:521-27; Sauer, 1993, *Methods Enzymol.*, 225:890-900) upstream of (*i.e.*, 5' to) the cell's endogenous genomic LGR8 polypeptide coding region. A plasmid containing a recombination site homologous to the site that was placed just upstream of the genomic LGR8 polypeptide coding region is introduced into the modified cell line along with the

appropriate recombinase enzyme. This recombinase causes the plasmid to integrate, via the plasmid's recombination site, into the recombination site located just upstream of the genomic LGR8 polypeptide coding region in the cell line (Baubonis and Sauer, 1993, *Nucleic Acids Res.* 21:2025-29; O'Gorman *et al.*, 1991, *Science* 251:1351-55). Any  
5 flanking sequences known to increase transcription (*e.g.*, enhancer/promoter, intron, translational enhancer), if properly positioned in this plasmid, would integrate in such a manner as to create a new or modified transcriptional unit resulting in *de novo* or increased LGR8 polypeptide production from the cell's endogenous LGR8 gene.

A further method to use the cell line in which the site specific recombination  
10 sequence had been placed just upstream of the cell's endogenous genomic LGR8 polypeptide coding region is to use homologous recombination to introduce a second recombination site elsewhere in the cell line's genome. The appropriate recombinase enzyme is then introduced into the two-recombination-site cell line, causing a recombination event (deletion, inversion, and translocation) (Sauer, 1994, *Curr. Opin.*  
15 *Biotechnol.*, 5:521-27; Sauer, 1993, *Methods Enzymol.*, 225:890-900) that would create a new or modified transcriptional unit resulting in *de novo* or increased LGR8 polypeptide production from the cell's endogenous LGR8 gene.

An additional approach for increasing, or causing, the expression of LGR8 polypeptide from a cell's endogenous LGR8 gene involves increasing, or causing, the  
20 expression of a gene or genes (*e.g.*, transcription factors) and/or decreasing the expression of a gene or genes (*e.g.*, transcriptional repressors) in a manner which results in *de novo* or increased LGR8 polypeptide production from the cell's endogenous LGR8 gene. This method includes the introduction of a non-naturally occurring polypeptide (*e.g.*, a polypeptide comprising a site specific DNA binding domain fused to a  
25 transcriptional factor domain) into the cell such that *de novo* or increased LGR8 polypeptide production from the cell's endogenous LGR8 gene results.

The present invention further relates to DNA constructs useful in the method of altering expression of a target gene. In certain embodiments, the exemplary DNA constructs comprise: (a) one or more targeting sequences, (b) a regulatory sequence, (c)  
30 an exon, and (d) an unpaired splice-donor site. The targeting sequence in the DNA construct directs the integration of elements (a) - (d) into a target gene in a cell such that

the elements (b) - (d) are operatively linked to sequences of the endogenous target gene. In another embodiment, the DNA constructs comprise: (a) one or more targeting sequences, (b) a regulatory sequence, (c) an exon, (d) a splice-donor site, (e) an intron, and (f) a splice-acceptor site, wherein the targeting sequence directs the integration of elements (a) - (f) such that the elements of (b) - (f) are operatively linked to the endogenous gene. The targeting sequence is homologous to the preselected site in the cellular chromosomal DNA with which homologous recombination is to occur. In the construct, the exon is generally 3' of the regulatory sequence and the splice-donor site is 3' of the exon.

If the sequence of a particular gene is known, such as the nucleic acid sequence of LGR8 polypeptide presented herein, a piece of DNA that is complementary to a selected region of the gene can be synthesized or otherwise obtained, such as by appropriate restriction of the native DNA at specific recognition sites bounding the region of interest. This piece serves as a targeting sequence upon insertion into the cell and will hybridize to its homologous region within the genome. If this hybridization occurs during DNA replication, this piece of DNA, and any additional sequence attached thereto, will act as an Okazaki fragment and will be incorporated into the newly synthesized daughter strand of DNA. The present invention, therefore, includes nucleotides encoding an LGR8 polypeptide, which nucleotides may be used as targeting sequences.

LGR8 polypeptide cell therapy, *e.g.*, the implantation of cells producing LGR8 polypeptides, is also contemplated. This embodiment involves implanting cells capable of synthesizing and secreting a biologically active form of LGR8 polypeptide. Such LGR8 polypeptide-producing cells can be cells that are natural producers of LGR8 polypeptides or may be recombinant cells whose ability to produce LGR8 polypeptides has been augmented by transformation with a gene encoding the desired LGR8 polypeptide or with a gene augmenting the expression of LGR8 polypeptide. Such a modification may be accomplished by means of a vector suitable for delivering the gene as well as promoting its expression and secretion. In order to minimize a potential immunological reaction in patients being administered an LGR8 polypeptide, as may occur with the administration of a polypeptide of a foreign species, it is preferred that the natural cells producing LGR8 polypeptide be of human origin and produce human LGR8

polypeptide. Likewise, it is preferred that the recombinant cells producing LGR8 polypeptide be transformed with an expression vector containing a gene encoding a human LGR8 polypeptide.

Implanted cells may be encapsulated to avoid the infiltration of surrounding tissue. Human or non-human animal cells may be implanted in patients in biocompatible, semipermeable polymeric enclosures or membranes that allow the release of LGR8 polypeptide, but that prevent the destruction of the cells by the patient's immune system or by other detrimental factors from the surrounding tissue. Alternatively, the patient's own cells, transformed to produce LGR8 polypeptides *ex vivo*, may be implanted directly into the patient without such encapsulation.

Techniques for the encapsulation of living cells are known in the art, and the preparation of the encapsulated cells and their implantation in patients may be routinely accomplished. For example, Baetge *et al.* (PCT Pub. No. WO 95/05452 and PCT/US94/09299) describe membrane capsules containing genetically engineered cells for the effective delivery of biologically active molecules. The capsules are biocompatible and are easily retrievable. The capsules encapsulate cells transfected with recombinant DNA molecules comprising DNA sequences coding for biologically active molecules operatively linked to promoters that are not subject to down-regulation *in vivo* upon implantation into a mammalian host. The devices provide for the delivery of the molecules from living cells to specific sites within a recipient. In addition, *see* U.S. Patent Nos. 4,892,538; 5,011,472; and 5,106,627. A system for encapsulating living cells is described in PCT Pub. No. WO 91/10425 (Aebischer *et al.*). *See also*, PCT Pub. No. WO 91/10470 (Aebischer *et al.*); Winn *et al.*, 1991, *Exper. Neurol.* 113:322-29; Aebischer *et al.*, 1991, *Exper. Neurol.* 111:269-75; and Tresco *et al.*, 1992, *ASAIO* 38:17-23.

*In vivo* and *in vitro* gene therapy delivery of LGR8 polypeptides is also envisioned. One example of a gene therapy technique is to use the LGR8 gene (either genomic DNA, cDNA, and/or synthetic DNA) encoding an LGR8 polypeptide that may be operably linked to a constitutive or inducible promoter to form a "gene therapy DNA construct." The promoter may be homologous or heterologous to the endogenous LGR8 gene, provided that it is active in the cell or tissue type into which the construct will be

inserted. Other components of the gene therapy DNA construct may optionally include DNA molecules designed for site-specific integration (*e.g.*, endogenous sequences useful for homologous recombination), tissue-specific promoters, enhancers or silencers, DNA molecules capable of providing a selective advantage over the parent cell, DNA molecules useful as labels to identify transformed cells, negative selection systems, cell specific binding agents (as, for example, for cell targeting), cell-specific internalization factors, transcription factors enhancing expression from a vector, and factors enabling vector production.

A gene therapy DNA construct can then be introduced into cells (either *ex vivo* or *in vivo*) using viral or non-viral vectors. One means for introducing the gene therapy DNA construct is by means of viral vectors as described herein. Certain vectors, such as retroviral vectors, will deliver the DNA construct to the chromosomal DNA of the cells, and the gene can integrate into the chromosomal DNA. Other vectors will function as episomes, and the gene therapy DNA construct will remain in the cytoplasm.

In yet other embodiments, regulatory elements can be included for the controlled expression of the LGR8 gene in the target cell. Such elements are turned on in response to an appropriate effector. In this way, a therapeutic polypeptide can be expressed when desired. One conventional control means involves the use of small molecule dimerizers or rapalogs to dimerize chimeric proteins which contain a small molecule-binding domain and a domain capable of initiating a biological process, such as a DNA-binding protein or transcriptional activation protein (*see* PCT Pub. Nos. WO 96/41865, WO 97/31898, and WO 97/31899). The dimerization of the proteins can be used to initiate transcription of the transgene.

An alternative regulation technology uses a method of storing proteins expressed from the gene of interest inside the cell as an aggregate or cluster. The gene of interest is expressed as a fusion protein that includes a conditional aggregation domain that results in the retention of the aggregated protein in the endoplasmic reticulum. The stored proteins are stable and inactive inside the cell. The proteins can be released, however, by administering a drug (*e.g.*, small molecule ligand) that removes the conditional aggregation domain and thereby specifically breaks apart the aggregates or clusters so



that the proteins may be secreted from the cell. See Aridor *et al.*, 2000, *Science* 287:816-17 and Rivera *et al.*, 2000, *Science* 287:826-30.

Other suitable control means or gene switches include, but are not limited to, the systems described herein. Mifepristone (RU486) is used as a progesterone antagonist.

5 The binding of a modified progesterone receptor ligand-binding domain to the progesterone antagonist activates transcription by forming a dimer of two transcription factors that then pass into the nucleus to bind DNA. The ligand-binding domain is modified to eliminate the ability of the receptor to bind to the natural ligand. The modified steroid hormone receptor system is further described in U.S. Patent No.  
10 5,364,791 and PCT Pub. Nos. WO 96/40911 and WO 97/10337.

Yet another control system uses ecdysone (a fruit fly steroid hormone) that binds to and activates an ecdysone receptor (cytoplasmic receptor). The receptor then translocates to the nucleus to bind a specific DNA response element (promoter from ecdysone-responsive gene). The ecdysone receptor includes a transactivation domain,  
15 DNA-binding domain, and ligand-binding domain to initiate transcription. The ecdysone system is further described in U.S. Patent No. 5,514,578 and PCT Pub. Nos. WO 97/38117, WO 96/37609, and WO 93/03162.

Another control means uses a positive tetracycline-controllable transactivator. This system involves a mutated tet repressor protein DNA-binding domain (mutated tet  
20 R-4 amino acid changes which resulted in a reverse tetracycline-regulated transactivator protein, *i.e.*, it binds to a tet operator in the presence of tetracycline) linked to a polypeptide which activates transcription. Such systems are described in U.S. Patent Nos. 5,464,758, 5,650,298, and 5,654,168.

Additional expression control systems and nucleic acid constructs are described in  
25 U.S. Patent Nos. 5,741,679 and 5,834,186, to Innovir Laboratories Inc.

*In vivo* gene therapy may be accomplished by introducing the gene encoding LGR8 polypeptide into cells via local injection of an LGR8 nucleic acid molecule or by other appropriate viral or non-viral delivery vectors. Hefti 1994, *Neurobiology* 25:1418-35. For example, a nucleic acid molecule encoding an LGR8 polypeptide may be  
30 contained in an adeno-associated virus (AAV) vector for delivery to the targeted cells (*see, e.g.*, Johnson, PCT Pub. No. WO 95/34670; PCT App. No. PCT/US95/07178). The

recombinant AAV genome typically contains AAV inverted terminal repeats flanking a DNA sequence encoding an LGR8 polypeptide operably linked to functional promoter and polyadenylation sequences.

Alternative suitable viral vectors include, but are not limited to, retrovirus, adenovirus, herpes simplex virus, lentivirus, hepatitis virus, parvovirus, papovavirus, poxvirus, alphavirus, coronavirus, rhabdovirus, paramyxovirus, and papilloma virus vectors. U.S. Patent No. 5,672,344 describes an *in vivo* viral-mediated gene transfer system involving a recombinant neurotrophic HSV-1 vector. U.S. Patent No. 5,399,346 provides examples of a process for providing a patient with a therapeutic protein by the delivery of human cells that have been treated *in vitro* to insert a DNA segment encoding a therapeutic protein. Additional methods and materials for the practice of gene therapy techniques are described in U.S. Patent Nos. 5,631,236 (involving adenoviral vectors), 5,672,510 (involving retroviral vectors), 5,635,399 (involving retroviral vectors expressing cytokines).

Nonviral delivery methods include, but are not limited to, liposome-mediated transfer, naked DNA delivery (direct injection), receptor-mediated transfer (ligand-DNA complex), electroporation, calcium phosphate precipitation, and microparticle bombardment (*e.g.*, gene gun). Gene therapy materials and methods may also include inducible promoters, tissue-specific enhancer-promoters, DNA sequences designed for site-specific integration, DNA sequences capable of providing a selective advantage over the parent cell, labels to identify transformed cells, negative selection systems and expression control systems (safety measures), cell-specific binding agents (for cell targeting), cell-specific internalization factors, and transcription factors to enhance expression by a vector as well as methods of vector manufacture. Such additional methods and materials for the practice of gene therapy techniques are described in U.S. Patent Nos. 4,970,154 (involving electroporation techniques), 5,679,559 (describing a lipoprotein-containing system for gene delivery), 5,676,954 (involving liposome carriers), 5,593,875 (describing methods for calcium phosphate transfection), and 4,945,050 (describing a process wherein biologically active particles are propelled at cells at a speed whereby the particles penetrate the surface of the cells and become incorporated into the interior of the cells), and PCT Pub. No. WO 96/40958 (involving nuclear ligands).

It is also contemplated that LGR8 gene therapy or cell therapy can further include the delivery of one or more additional polypeptide(s) in the same or a different cell(s). Such cells may be separately introduced into the patient, or the cells may be contained in a single implantable device, such as the encapsulating membrane described above, or the cells may be separately modified by means of viral vectors.

A means to increase endogenous LGR8 polypeptide expression in a cell via gene therapy is to insert one or more enhancer elements into the LGR8 polypeptide promoter, where the enhancer elements can serve to increase transcriptional activity of the LGR8 gene. The enhancer elements used will be selected based on the tissue in which one desires to activate the gene – enhancer elements known to confer promoter activation in that tissue will be selected. For example, if a gene encoding an LGR8 polypeptide is to be “turned on” in T-cells, the *lck* promoter enhancer element may be used. Here, the functional portion of the transcriptional element to be added may be inserted into a fragment of DNA containing the LGR8 polypeptide promoter (and optionally, inserted into a vector and/or 5’ and/or 3’ flanking sequences) using standard cloning techniques. This construct, known as a “homologous recombination construct,” can then be introduced into the desired cells either *ex vivo* or *in vivo*.

Gene therapy also can be used to decrease LGR8 polypeptide expression by modifying the nucleotide sequence of the endogenous promoter. Such modification is typically accomplished via homologous recombination methods. For example, a DNA molecule containing all or a portion of the promoter of the LGR8 gene selected for inactivation can be engineered to remove and/or replace pieces of the promoter that regulate transcription. For example, the TATA box and/or the binding site of a transcriptional activator of the promoter may be deleted using standard molecular biology techniques; such deletion can inhibit promoter activity thereby repressing the transcription of the corresponding LGR8 gene. The deletion of the TATA box or the transcription activator binding site in the promoter may be accomplished by generating a DNA construct comprising all or the relevant portion of the LGR8 polypeptide promoter (from the same or a related species as the LGR8 gene to be regulated) in which one or more of the TATA box and/or transcriptional activator binding site nucleotides are mutated via substitution, deletion and/or insertion of one or more nucleotides. As a

result, the TATA box and/or activator binding site has decreased activity or is rendered completely inactive. This construct, which also will typically contain at least about 500 bases of DNA that correspond to the native (endogenous) 5' and 3' DNA sequences adjacent to the promoter segment that has been modified, may be introduced into the appropriate cells (either *ex vivo* or *in vivo*) either directly or via a viral vector as described herein. Typically, the integration of the construct into the genomic DNA of the cells will be via homologous recombination, where the 5' and 3' DNA sequences in the promoter construct can serve to help integrate the modified promoter region via hybridization to the endogenous chromosomal DNA.

### Therapeutic Uses

LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof can be used to treat, diagnose, ameliorate, or prevent a number of diseases, disorders, or conditions, including those recited herein.

LGR8 polypeptide agonists and antagonists include those molecules which regulate LGR8 polypeptide activity and either increase or decrease at least one activity of the mature form of the LGR8 polypeptide. Agonists or antagonists may be co-factors, such as a protein, peptide, carbohydrate, lipid, or small molecular weight molecule, which interact with LGR8 polypeptide and thereby regulate its activity. Potential polypeptide agonists or antagonists include antibodies that react with either soluble or membrane-bound forms of LGR8 polypeptides that comprise part or all of the extracellular domains of the said proteins. Molecules that regulate LGR8 polypeptide expression typically include nucleic acids encoding LGR8 polypeptide that can act as anti-sense regulators of expression.

Seven members of the glycoprotein hormone receptor subfamily have been previously identified. Among these are thyroid stimulating hormone (TSH) receptor, follicle stimulating hormone (FSH) receptor, luteinizing hormone (LH)/chorionic gonadotropin (CG) receptor, and leucine-rich repeat-containing G-protein coupled receptor (LGR) 4, LGR5, LGR6, and LGR7. The TSH, FSH, and LH/CG receptors have been well characterized functionally, with signaling through these receptors playing an important role in the proliferation and differentiation of the thyroid gland and the gonads.

Signaling through the TSH receptor is known to influence basal metabolism by regulating the production of thyroid hormones. Autoimmune antibodies against epitopes of the N-terminal extracellular domain of the TSH receptor cause various thyroid hormone-associated metabolic disorders. Such autoimmune antibodies can be antagonistic and cause a form of hypothyroidism (a subgroup of Hashimoto's thyroiditis) or can be agonistic and cause a form of hyperthyroidism (Graves' disease). Signaling through the FSH and LH/CG receptors is known to play a critical role in the maintenance of reproductive function in males and females (*i.e.*, gonadal maturation and gonadal steroid production). Additionally, signaling through the LH/CG receptor is known to play an important role in the maintenance of pregnancy by stimulating the corpus luteum to produce steroid hormones during the first trimester. Because the TSH, FSH, and LH/CG receptors are known to have important developmental (*i.e.*, proliferation and differentiation) and physiological functions, it is likely that LGR8 also plays an important role in development and in human physiology.

Since LGR8 polypeptide expression has been detected in skeletal muscle, LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting skeletal muscle. Examples of such diseases and conditions include, but are not limited to, cachexia and muscular dystrophy. Other diseases and conditions associated with skeletal muscle development and function are encompassed within the scope of this invention.

Since LGR8 polypeptide expression has been detected in the uterus, LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting skeletal muscle. Examples of such diseases and conditions include, but are not limited to, miscarriage, endometriosis, uterine cancer, and female infertility. Other diseases and conditions associated with uterine development and function are encompassed within the scope of this invention.

Since LGR8 polypeptide expression has been detected in the adrenal gland, LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting the adrenal gland. Examples of such diseases and conditions include, but are not limited to, Cushing's disease and

Addison's disease. Other diseases and conditions associated with the development and function of the adrenal gland are encompassed within the scope of this invention.

Since LGR8 polypeptide expression has been detected in the testes, LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in  
5 diagnosing or treating diseases and conditions affecting the testes. Examples of such diseases and conditions include, but are not limited to, male infertility and testicular carcinoma. Other diseases and conditions associated with the development and function of the testes are encompassed within the scope of this invention.

Since LGR8 polypeptide expression has been detected in the bone marrow, LGR8  
10 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting the bone marrow. Examples of such diseases and conditions include, but are not limited to, leukemia. Other diseases and conditions associated with the development and function of the bone marrow are encompassed within the scope of this invention.

Since LGR8 polypeptide expression has been detected in the fetal kidney, LGR8  
15 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting the kidney. Examples of such diseases and conditions include, but are not limited to, anemia, hypertension, and low blood pressure. Other diseases and conditions associated with the development and  
20 function of the kidney are encompassed within the scope of this invention.

Since LGR8 polypeptide expression has been detected in the fetal ovary, LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting the ovaries. Examples of such  
25 diseases and conditions include, but are not limited to, female infertility and ovarian cancer. Other diseases and conditions associated with the development and function of the ovaries are encompassed within the scope of this invention.

Since LGR8 polypeptide is likely to play a role in cell proliferation and differentiation, LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions that modulate cell  
30 proliferation and differentiation. For example, the LGR8 molecules of the invention may be used to increase cell proliferation and differentiation. Examples of diseases and

conditions that may be treated by increasing cell proliferation and differentiation include, but are not limited to, tissue damage and degeneration (such as that caused by cancer therapy, infections, autoimmune diseases, or disorders), aging, and wound healing. Other diseases and conditions that could be treated by increasing cell proliferation and differentiation are encompassed within the scope of this invention.

Alternatively, the LGR8 molecules of the invention may be used to decrease cell proliferation and differentiation. Examples of diseases and conditions that may be treated by decreasing cell proliferation and differentiation include, but are not limited to, cancer, hyperplasia, and hypertrophy. Other diseases and conditions that could be treated by decreasing cell proliferation and differentiation are encompassed within the scope of this invention.

Agonists or antagonists of LGR8 polypeptide function may be used (simultaneously or sequentially) in combination with one or more cytokines, growth factors, antibiotics, anti-inflammatories, and/or chemotherapeutic agents as is appropriate for the condition being treated.

Other diseases or disorders caused by or mediated by undesirable levels of LGR8 polypeptides are encompassed within the scope of the invention. Undesirable levels include excessive levels of LGR8 polypeptides and sub-normal levels of LGR8 polypeptides.

#### Uses of LGR8 Nucleic Acids and Polypeptides

Nucleic acid molecules of the invention (including those that do not themselves encode biologically active polypeptides) may be used to map the locations of the LGR8 gene and related genes on chromosomes. Mapping may be done by techniques known in the art, such as PCR amplification and *in situ* hybridization.

LGR8 nucleic acid molecules (including those that do not themselves encode biologically active polypeptides), may be useful as hybridization probes in diagnostic assays to test, either qualitatively or quantitatively, for the presence of an LGR8 nucleic acid molecule in mammalian tissue or bodily fluid samples.

Other methods may also be employed where it is desirable to inhibit the activity of one or more LGR8 polypeptides. Such inhibition may be effected by nucleic acid

molecules that are complementary to and hybridize to expression control sequences (triple helix formation) or to LGR8 mRNA. For example, antisense DNA or RNA molecules, which have a sequence that is complementary to at least a portion of an LGR8 gene can be introduced into the cell. Anti-sense probes may be designed by available techniques using the sequence of the LGR8 gene disclosed herein. Typically, each such antisense molecule will be complementary to the start site (5' end) of each selected LGR8 gene. When the antisense molecule then hybridizes to the corresponding LGR8 mRNA, translation of this mRNA is prevented or reduced. Anti-sense inhibitors provide information relating to the decrease or absence of an LGR8 polypeptide in a cell or organism.

Alternatively, gene therapy may be employed to create a dominant-negative inhibitor of one or more LGR8 polypeptides. In this situation, the DNA encoding a mutant polypeptide of each selected LGR8 polypeptide can be prepared and introduced into the cells of a patient using either viral or non-viral methods as described herein. Each such mutant is typically designed to compete with endogenous polypeptide in its biological role.

In addition, an LGR8 polypeptide, whether biologically active or not, may be used as an immunogen, that is, the polypeptide contains at least one epitope to which antibodies may be raised. Selective binding agents that bind to an LGR8 polypeptide (as described herein) may be used for *in vivo* and *in vitro* diagnostic purposes, including, but not limited to, use in labeled form to detect the presence of LGR8 polypeptide in a body fluid or cell sample. The antibodies may also be used to prevent, treat, or diagnose a number of diseases and disorders, including those recited herein. The antibodies may bind to an LGR8 polypeptide so as to diminish or block at least one activity characteristic of an LGR8 polypeptide, or may bind to a polypeptide to increase at least one activity characteristic of an LGR8 polypeptide (including by increasing the pharmacokinetics of the LGR8 polypeptide).

LGR8 polypeptides can be used to clone LGR8 ligands using an "expression cloning" strategy. Radiolabeled (<sup>125</sup>Iodine) LGR8 polypeptide or "affinity/activity-tagged" LGR8 polypeptide (such as an Fc fusion or an alkaline phosphatase fusion) can be used in binding assays to identify a cell type, cell line, or tissue that expresses an



LGR8 ligand. RNA isolated from such cells or tissues can then be converted to cDNA, cloned into a mammalian expression vector, and transfected into mammalian cells (e.g., COS or 293) to create an expression library. Radiolabeled or tagged LGR8 polypeptide can then be used as an affinity reagent to identify and isolate the subset of cells in this library expressing an LGR8 ligand. DNA is then isolated from these cells and transfected into mammalian cells to create a secondary expression library in which the fraction of cells expressing the LGR8 ligand would be many-fold higher than in the original library. This enrichment process can be repeated iteratively until a single recombinant clone containing the LGR8 ligand is isolated. Isolation of LGR8 ligands is useful for identifying or developing novel agonists and antagonists of the LGR8 signaling pathway. Such agonists and antagonists include LGR8 ligands, anti-LGR8 ligand antibodies, small molecules or antisense oligonucleotides.

The murine and human LGR8 nucleic acids of the present invention are also useful tools for isolating the corresponding chromosomal LGR8 polypeptide genes. For example, mouse chromosomal DNA containing LGR8 sequences can be used to construct knockout mice, thereby permitting an examination of the *in vivo* role for LGR8 polypeptide. The human LGR8 genomic DNA can be used to identify heritable tissue-degenerating diseases.

The following examples are intended for illustration purposes only, and should not be construed as limiting the scope of the invention in any way.

#### Example 1: Cloning of the Human and Murine LGR8 Genes

Generally, materials and methods as described in Sambrook *et al. supra* were used to clone and analyze the gene encoding murine LGR8 polypeptide.

An informatics-based search of human genomics sequences led to the identification of the 3' end of the nucleotide sequence encoding human LGR8. A similar search led to the identification of the 5' end of the nucleotide sequence encoding the mature form (*i.e.*, lacking the signal peptide) of human LGR8. These sequences were then used to design gene specific oligonucleotides for the identification of cDNA sources and the generation of cDNA clones, using various PCR strategies. Several highly homologous, but not identical, LGR8 sequences were thus isolated. An analysis of these

sequences led to the identification of four nucleotide sequences encoding the mature forms (*i.e.*, the forms lacking an LGR8 start codon and the nucleotide sequence encoding the signal peptide) of LGR8-A, LGR8-B, LGR8-C, and LGR8-D.

5 The nucleotide sequence encoding the mature form of human LGR8-A was obtained in amplification reactions using 5 µl of a human adrenal Marathon Ready cDNA template (Clontech Laboratories; Palo Alto, CA), 1.0 µm each of the amplimers 5'-T-G-C-C-A-A-A-A-G-G-A-T-A-T-T-T-T-C-C-C-T-G-T-G-G-G-A-A-T-C-T-T-A-3' (SEQ ID NO: 27) and 5'-C-T-A-G-G-A-AA-C-T-G-G-T-T-T-C-A-T-T-A-T-A-C-T-G-T-C-T-C-C-A-A-G-T-G-T-T-A-T-T-T-T-G-T-T-C-A-3' (SEQ ID NO: 28), 200 µm of dNTPs, 10 2.5 U of *Pfu*Turbo DNA polymerase (Stratagene; La Jolla, CA), and 5 µL 10X *Pfu*Turbo DNA polymerase reaction buffer in a final volume of 50 µL. Reactions were performed at 94°C for 1 minute for one cycle; 94°C for 10 seconds, 60°C for 20 seconds, and 72°C for 5 minutes for 45 cycles; and 72°C for 7 minutes for one cycle.

15 The amplification mixture was separated on an agarose gel, the PCR products were isolated from the gel, and the products were then blunt-end cloned into pPCR-Script Amp SK(+) (Stratagene). A number of clones were sequenced, each containing highly homologous (but not identical) LGR8 nucleotide sequences. These sequences were used to compile the four nucleotide sequences encoding the mature forms (*i.e.*, the form lacking a start codon and a signal peptide) of LGR8-A, LGR8-B, LGR8-C, and LGR8-D.

20 To isolate cDNA sequences corresponding to the 5' end of the cDNA sequence for the immature form of human LGR8, 5' RACE was performed using 5 µL of a human adrenal Marathon Ready cDNA template, 1.0 µm each of the primers 5'-C-C-A-T-C-C-T-A-A-T-A-C-G-A-C-T-C-A-C-T-A-T-A-G-G-G-C-3' (SEQ ID NO: 29) and 5'-A-T-T-G-T-C-A-T-C-T-A-G-A-A-T-T-A-G-C-C-A-A-G-T-T-A-G-C-T-G-A-T-3' (SEQ ID NO: 25 30), 200 µm of dNTPs, 1 µL of 50X Advantage2 Polymerase Mix (Clontech Laboratories), and 5 µL 10X Advantage2 PCR buffer in a final volume of 50 µL. Reactions were performed at 94°C for 1 minute for one cycle; 94°C for 20 seconds and 68°C for 3 minutes for 35 cycles; and 68°C for 3 minutes for one cycle. Nested PCR was performed using 0.1 µL of the 5' RACE amplification product, 1.0 µm each of the 30 primers 5'-A-C-T-C-A-C-T-A-T-A-G-G-G-C-T-C-G-A-G-C-G-G-C-3' (SEQ ID NO: 32) and 5'-A-T-A-T-T-C-C-A-G-G-T-C-T-G-A-G-G-G-T-T-G-T-G-A-T-3' (SEQ ID

NO: 33), 200 µm of dNTPs, 1 µL of 50X Advantage2 Polymerase Mix, and 5 µL 10X Advantage2 PCR buffer in a final volume of 50 µL. Nested PCR Reactions were performed at 94°C for 1 minute for one cycle; 94°C for 20 seconds and 72°C for 3 minutes for 28 cycles; and 72°C for 3 minutes for one cycle.

5 Alternatively, 5' RACE was performed using 5 µL of a human adrenal Marathon Ready cDNA template, 1.0 µm each of the primers 5'-C-C-A-T-C-C-T-A-A-T-A-C-G-A-C-T-C-A-C-T-A-T-A-G-G-G-C-3' (SEQ ID NO: 29) and 5'-A-A-C-A-A-G-G-A-A-T-T-A-A-T-C-C-C-G-T-A-A-A-C-A-A-G-3' (SEQ ID NO: 31), 200 µm of dNTPs, 1 µL of 50X Advantage2 Polymerase Mix, and 5 µL 10X Advantage2 PCR buffer in a final  
10 volume of 50 µL. Reactions were performed at 94°C for 1 minute for one cycle; 94°C for 20 seconds and 68°C for 3 minutes for 35 cycles; and 68°C for 3 minutes for one cycle. Nested PCR was performed using 0.1 µL of the 5' RACE amplification product, 1.0 µm each of the primers 5'-A-C-T-C-A-C-T-A-T-A-G-G-G-C-T-C-G-A-G-C-G-G-C-3' (SEQ ID NO: 32) and 5'-A-T-A-T-T-C-C-A-G-G-T-C-T-G-A-G-G-G-T-T-G-T-G-A-T-3'  
15 (SEQ ID NO: 34), 200 µm of dNTPs, 1 µL of 50X Advantage2 Polymerase Mix, and 5 µL 10X Advantage2 PCR buffer in a final volume of 50 µL. Nested PCR Reactions were performed at 94°C for 1 minute for one cycle; 94°C for 20 seconds and 72°C for 3 minutes for 28 cycles; and 72°C for 3 minutes for one cycle.

The PCR products generated in the nested amplification reactions were separated  
20 on an agarose gel, the PCR products were isolated from the gel, and the products were then cloned into pCR2.1 (Invitrogen; Carlsbad, CA). A number of clones were sequenced, each containing sequences that were homologous to those encoding the mature form of LGR8, but also encoding a methionine and full signal peptide. These sequences were used to compile the four nucleotide sequences encoding the full-length  
25 cDNA (*i.e.*, that encoding a start codon and full signal peptide) for LGR8-A, LGR8-B, LGR8-C, and LGR8-D.

The LGR8-A coding sequence encodes a large N-terminal leucine-rich repeat-containing extracellular domain, seven predicted transmembrane domains, and a cytoplasmic C-terminal region. Sequence analysis of the full-length LGR8-A coding  
30 sequence indicated that the cDNA comprises a 2262 bp open reading frame encoding a protein of 754 amino acids (Figures 1A-1D). The mature form of LGR8-A is 718 amino

acids in length. LGR8-A is most closely related to glycoprotein hormone receptor LGR7 (Figures 10A-10B).

The LGR8-B coding sequence is identical to the LGR8-A coding sequence with the exception that the LGR8-B coding sequence lacks a portion of the sequence encoding the N-terminal extracellular domain. Sequence analysis of the full-length LGR8-B coding sequence indicated that the cDNA comprises a 2190 bp open reading frame encoding a protein of 730 amino acids (Figures 3A-3D). The mature form of LGR8-B is 694 amino acids in length.

The LGR8-C coding sequence is identical to the LGR8-A coding sequence with the exception that the LGR8-C coding sequence lacks a portion of sequence encoding the N-terminal extracellular domain. Sequence analysis of the full-length LGR8-C coding sequence indicated that the cDNA comprises a 2046 bp open reading frame encoding a protein of 682 amino acids (Figures 5A-5D). The mature form of LGR8-C is 646 amino acids in length.

The LGR8-D coding sequence consists of sequence encoding approximately 90% of the N-terminal extracellular domain of the LGR8-B coding sequence, but lacks sequence encoding the transmembrane domains and the cytoplasmic C-terminal region. Sequence analysis of the full-length LGR8-D coding sequence indicated that the cDNA comprises a 1098 bp open reading frame encoding a protein of 366 amino acids (Figures 7A-7B). The mature form of LGR8-D is 330 amino acids in length.

To identify cDNA sequences encoding murine LGR8-A, homology-based BLAST searches of a human genomic database were performed using the amino acid sequence of human LGR8-A. A number of sequences sharing a high degree of homology were found within a 213 kb mouse public genomic sequence (Accession No. AC077689). No exons, genes, or homologies to known genes were identified in the AC077689 sequence record. The sequences thus identified were curated by hand and electronically compiled to yield the complete nucleotide sequence encoding murine LGR8-A (Figures 8A-8D). A sequence comparison of the mature human and murine LGR8-A sequences indicates that the sequences share an 86.6% similarity and an 83.1% identity (Figures 11A-11B). A sequence comparison of the N-terminal extracellular domains (absent the signal peptide) of human and murine LGR8-A indicates that the sequences share an

85.5% similarity and an 82.3% identity (Figure 12).

#### Example 2: LGR8 mRNA Expression

Since it was not possible to obtain a hybridization signal on various human  
5 multiple tissue Northern blots (Clontech) using a PCR fragment generated from the  
human LGR8 coding sequence as a probe, the expression of LGR8 mRNA was analyzed  
by PCR.

Intron-spanning PCR was first performed on human Marathon Ready cDNAs  
(Clontech) for fetal adrenal, brain, kidney, liver, lung, spleen, thymus, and adult bone  
10 marrow, heart, kidney, liver, lung, lymph node, pancreas, placenta, retina, skeletal  
muscle, small intestine, spleen, testis, thymus, pituitary, adrenal, and prostate.

The human cDNA libraries to be tested in PCR were prepared as follows. Total  
RNA was extracted from the appropriate tissue or cell line using standard RNA  
extraction procedures and poly-A<sup>+</sup> RNA was selected from this total RNA using standard  
15 procedures. Random-primed or oligo-dT primed cDNA was synthesized from this poly-  
A<sup>+</sup> RNA using the Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning  
kit (Gibco-BRL), according to the manufacturer's suggested protocols, or other suitable  
procedures known to those skilled in the art. The resulting cDNA was digested with  
suitable restriction endonucleases and was then ligated into pSPORT-1, or other suitable  
20 vector known to those skilled in the art. Ligation products were transformed into *E. coli*  
using standard techniques, and bacterial transformants were selected on culture plates  
containing ampicillin, tetracycline, kanamycin, or chloramphenicol. The cDNA library  
consisted of all, or a subset, of these transformants.

The expression of LGR8 mRNA in these samples was analyzed in amplification  
25 reactions using 3 µL of human Marathon Ready cDNA as a template, 1.0 µm each of the  
amplimers 5'-C-T-G-C-T-T-G-G-A-A-A-T-C-T-T-T-T-G-T-C-A-3' (SEQ ID NO:  
35) and 5'-T-T-T-T-C-C-A-G-G-T-C-G-A-A-T-G-T-T-A-C-T-G-A-3' (SEQ ID NO: 36),  
200 µm of dNTPs, 2.5 U of Taq polymerase (Boehringer Mannheim; Indianapolis, IN),  
and 2.5 µL 10X PCR-reaction buffer (Boehringer Mannheim) in a final volume of 25 µL.  
30 Reactions were performed at 94°C for 3 minutes for one cycle; 94°C for 30 seconds and  
70°C for 1.5 minutes for 35 cycles; and 70°C for 10 minutes for one cycle. The

amplification mixtures were separated on an agarose gel, and PCR products of the expected size (319 bp) were identified in adult skeletal muscle, testis, and adrenal gland.

Intron-spanning PCR was next performed on proprietary oligo-dT primed and random primed human cDNA libraries for the following tissues: fetal stomach (oligo-dT primed), fetal stomach (random primed), pons/medulla (oligo-dT primed), breast tumor T1485 (oligo-dT primed), breast tumor T1485 (random primed), ovary tumor T22 (oligo-dT primed), ovary tumor T22 (random primed), fetal thymus (oligo-dT primed), fetal thymus (random primed), fetal mesentery (oligo-dT primed), fetal mesentery (random primed), placenta (oligo-dT primed), placenta (random primed), Multiple cell lines [A204, A673, Hs729T, HISM and RD (oligo-dT primed)], Multiple cell lines [A204, A673, Hs729T, HISM and RD (random primed)], fetal pancreas (oligo-dT primed), fetal pancreas (random primed), lymphoma cell lines (oligo-dT primed), lymphoma cell lines (random primed), ovary tumor T23 (oligo-dT primed), ovary tumor T23 (random primed), colon tumor T25 (oligo-dT primed), colon tumor T25 (random primed), adult T-cells (oligo-dT primed), normalized fetal tissue (random primed), fetal heart (oligo-dT primed), fetal heart (random primed), fetal bladder (oligo-dT primed), fetal bladder (random primed), fetal kidney (oligo-dT primed), fetal kidney (random primed), lung tumor T27 (oligo-dT primed), lung tumor T27 (random primed), fetal liver (oligo-dT primed), cytoplasmic breast carcinoma cell lines (oligo-dT primed), cytoplasmic breast carcinoma cell lines (random primed), fetal spleen (oligo-dT primed), fetal spleen (random primed), uterus (oligo-dT primed), uterus (random primed), adrenal (oligo-dT primed), adrenal (random primed), forebrain (oligo-dT primed), forebrain (random primed), testis (oligo-dT primed), testis (random primed), colon tumor T24 (oligo-dT primed), colon tumor T24 (random primed), fetal heart (oligo-dT primed), fetal scalp (oligo-dT primed), fetal scalp (random primed), fetal lung (oligo-dT primed), fetal lung (random primed), trachea (oligo-dT primed), trachea (random primed), cerebellum (oligo-dT primed), midbrain LNV block 10 (oligo-dT primed), midbrain LNV block 10 (random primed), prostate tumor T1940 (random primed), fetal ovary (oligo-dT primed), fetal calveria (oligo-dT primed), fetal calveria (random primed), fetal gall bladder (oligo-dT primed), fetal gall bladder (random primed), spinal column (oligo-dT primed), spinal column (random primed), thalamus (oligo-dT primed), prostate tumor T1175 (oligo-dT

primed), prostate tumor T1175 (random primed), breast tumor T1543 (oligo-dT primed), breast tumor T1543 (random primed), fetal skin (oligo-dT primed), fetal skin (random primed), fetal femur (oligo-dT primed), fetal femur (random primed), T-lymphocytes (oligo-dT primed), T-lymphocytes (random primed), limb bone (oligo-dT primed), and limb bone (random primed). These cDNA libraries were prepared as described above.

The expression of LGR8 mRNA in these cDNA samples was analyzed essentially as described above using 50 ng of cDNA as a template. The amplification mixtures were separated on an agarose gel, and PCR products of the expected size were identified in fetal kidney, ovary, and femur, and adult uterus, adrenal gland, and forebrain.

Quantitative PCR using the PRISM Taqman system was performed on human cDNA for brain, heart, cerebellum, spleen, lung, skeletal muscle, kidney, testes, small intestine, pancreas, bone marrow, hippocampus, thalamus, spinal cord, uterus, prostate, stomach, pituitary, adrenal gland, thyroid, salivary gland, mammary gland, liver, and HeLa cells (mRNA obtained from Clontech Laboratories; cDNA prepared using the SuperScript Amplification System; Gibco BRL). Human cDNA for A432 cells, 293 cells, and white adipose tissue was also tested (mRNA extracted from the tissue or cell line using standard RNA extraction procedures; cDNA prepared using the SuperScript Amplification System).

PRISM Taqman reactions to assess LGR8 mRNA expression were performed using 50 ng of cDNA as a template, 300 nm each of the primers 5'-A-T-G-C-C-T-T-G-C-T-G-T-G-G-A-T-G-G-A-G-3' (SEQ ID NO: 37) and 5'-A-C-T-T-C-G-G-T-G-G-A-C-A-G-C-A-T-G-G-3' (SEQ ID NO: 38), 200 nm of the flurogenic probe 5'-(6-FAM)-C-G-T-G-C-A-G-T-G-C-C-G-C-C-T-C-A-T-G-G-(TAMRA)-3' (Primer Express, PE BioSystems, Foster City, CA; SEQ ID NO: 39; wherein "6-FAM" is the 5' reporter dye 6-carboxy-fluorescein and "TAMRA" is the 3' quencher 6-carboxytetramethylrhodamine), 200  $\mu$ m of dNTPs, 0.5 U of Uracil-N-glycosylase (AmpErase UNG; PE BioSystems), 1.25 U AmpliTaq Gold DNA polymerase (PE BioSystems), and 5  $\mu$ L 10X Taqman reaction buffer (PE BioSystems) in a final volume of 50  $\mu$ L. Reactions were performed at 50°C for 2 minutes and 95°C for 10 minutes for one cycle and 95°C for 15 seconds and 60°C for 1 minute for 40 cycles. LGR8 mRNA expression was normalized against the human housekeeper protein cyclophilin by measuring cyclophilin

mRNA expression in each of the cDNA samples analyzed for LGR8 mRNA expression.

PRISM Taqman reactions to assess cyclophilin mRNA expression were performed as described above using 300 nm each of the primers 5'-G-T-C-G-A-C-G-G-C-G-A-G-C-C-C-3' (SEQ ID NO: 40) and 5'-T-C-T-T-T-G-G-G-A-C-C-T-T-G-T-C-T-G-C-A-A-3' (SEQ ID NO: 41) and 200 nm of the flurogenic probe 5'-(6-FAM)-T-G-G-G-C-C-G-C-G-T-C-T-C-C-T-T-T-G-A-G-C-T-(TAMRA)-3' (Primer Express, PE BioSystems, Foster City, CA; SEQ ID NO: 42; wherein "6-FAM" is the 5' reporter dye 6-carboxy-fluorescein and "TAMRA" is the 3' quencher 6-carboxytetramethylrhodamine). The highest levels of LGR8 mRNA expression were detected in skeletal muscle and uterus. Lower levels were found in adrenal and testis, with lower levels still in thalamus and bone marrow.

The expression of LGR8 mRNA is localized by *in situ* hybridization. A panel of normal embryonic and adult mouse tissues is fixed in 4% paraformaldehyde, embedded in paraffin, and sectioned at 5  $\mu$ m. Sectioned tissues are permeabilized in 0.2 M HCl, digested with Proteinase K, and acetylated with triethanolamine and acetic anhydride. Sections are prehybridized for 1 hour at 60°C in hybridization solution (300 mM NaCl, 20 mM Tris-HCl, pH 8.0, 5 mM EDTA, 1X Denhardt's solution, 0.2% SDS, 10 mM DTT, 0.25 mg/ml tRNA, 25  $\mu$ g/ml polyA, 25  $\mu$ g/ml polyC and 50% formamide) and then hybridized overnight at 60°C in the same solution containing 10% dextran and  $2 \times 10^4$  cpm/ $\mu$ l of a  $^{33}$ P-labeled antisense riboprobe complementary to the human LGR8 gene. The riboprobe is obtained by *in vitro* transcription of a clone containing human LGR8 cDNA sequences using standard techniques.

Following hybridization, sections are rinsed in hybridization solution, treated with RNaseA to digest unhybridized probe, and then washed in 0.1X SSC at 55°C for 30 minutes. Sections are then immersed in NTB-2 emulsion (Kodak, Rochester, NY), exposed for 3 weeks at 4°C, developed, and counterstained with hematoxylin and eosin. Tissue morphology and hybridization signal are simultaneously analyzed by darkfield and standard illumination for brain (one sagittal and two coronal sections), gastrointestinal tract (esophagus, stomach, duodenum, jejunum, ileum, proximal colon, and distal colon), pituitary, liver, lung, heart, spleen, thymus, lymph nodes, kidney,



adrenal, bladder, pancreas, salivary gland, male and female reproductive organs (ovary, oviduct, and uterus in the female; and testis, epididymus, prostate, seminal vesicle, and vas deferens in the male), BAT and WAT (subcutaneous, peri-renal), bone (femur), skin, breast, and skeletal muscle.

5

### Example 3: Production of LGR8 Polypeptides

#### A. Expression of LGR8 Polypeptides in Bacteria

PCR is used to amplify template DNA sequences encoding an LGR8 polypeptide using primers corresponding to the 5' and 3' ends of the sequence. The amplified DNA products may be modified to contain restriction enzyme sites to allow for insertion into expression vectors. PCR products are gel purified and inserted into expression vectors using standard recombinant DNA methodology. An exemplary vector, such as pAMG21 (ATCC no. 98113) containing the lux promoter and a gene encoding kanamycin resistance is digested with Bam HI and Nde I for directional cloning of inserted DNA. The ligated mixture is transformed into an *E. coli* host strain by electroporation and transformants are selected for kanamycin resistance. Plasmid DNA from selected colonies is isolated and subjected to DNA sequencing to confirm the presence of the insert.

Transformed host cells are incubated in 2xYT medium containing 30 µg/mL kanamycin at 30°C prior to induction. Gene expression is induced by the addition of N-(3-oxohexanoyl)-dl-homoserine lactone to a final concentration of 30 ng/mL followed by incubation at either 30°C or 37°C for six hours. The expression of LGR8 polypeptide is evaluated by centrifugation of the culture, resuspension and lysis of the bacterial pellets, and analysis of host cell proteins by SDS-polyacrylamide gel electrophoresis.

Inclusion bodies containing LGR8 polypeptide are purified as follows. Bacterial cells are pelleted by centrifugation and resuspended in water. The cell suspension is lysed by sonication and pelleted by centrifugation at 195,000 xg for 5 to 10 minutes. The supernatant is discarded, and the pellet is washed and transferred to a homogenizer. The pellet is homogenized in 5 mL of a Percoll solution (75% liquid Percoll and 0.15 M NaCl) until uniformly suspended and then diluted and centrifuged at 21,600 xg for 30

minutes. Gradient fractions containing the inclusion bodies are recovered and pooled. The isolated inclusion bodies are analyzed by SDS-PAGE.

A single band on an SDS polyacrylamide gel corresponding to *E. coli*-produced LGR8 polypeptide is excised from the gel, and the N-terminal amino acid sequence is  
5 determined essentially as described by Matsudaira *et al.*, 1987, *J. Biol. Chem.* 262:10-35.

#### B. Expression of LGR8 Polypeptide in Mammalian Cells

PCR is used to amplify template DNA sequences encoding an LGR8 polypeptide using primers corresponding to the 5' and 3' ends of the sequence. The amplified DNA  
10 products may be modified to contain restriction enzyme sites to allow for insertion into expression vectors. PCR products are gel purified and inserted into expression vectors using standard recombinant DNA methodology. An exemplary expression vector, pCEP4 (Invitrogen, Carlsbad, CA), that contains an Epstein-Barr virus origin of replication, may be used for the expression of LGR8 polypeptides in 293-EBNA-1 cells.  
15 Amplified and gel purified PCR products are ligated into pCEP4 vector and introduced into 293-EBNA cells by lipofection. The transfected cells are selected in 100 µg/mL hygromycin and the resulting drug-resistant cultures are grown to confluence. The cells are then cultured in serum-free media for 72 hours. The conditioned media is removed and LGR8 polypeptide expression is analyzed by SDS-PAGE.

20 LGR8 polypeptide expression may be detected by silver staining. Alternatively, LGR8 polypeptide is produced as a fusion protein with an epitope tag, such as an IgG constant domain or a FLAG epitope, which may be detected by Western blot analysis using antibodies to the peptide tag.

LGR8 polypeptides may be excised from an SDS-polyacrylamide gel, or LGR8  
25 fusion proteins are purified by affinity chromatography to the epitope tag, and subjected to N-terminal amino acid sequence analysis as described herein.

#### C. Purification of LGR8 Polypeptide from Mammalian Cells

LGR8 polypeptide expression constructs are introduced into 293 EBNA or CHO  
30 cells using either a lipofection or calcium phosphate protocol.

To conduct functional studies on the LGR8 polypeptides that are produced, large quantities of conditioned media are generated from a pool of hygromycin selected 293 EBNA clones. The cells are cultured in 500 cm Nunc Triple Flasks to 80% confluence before switching to serum-free media a week prior to harvesting the media. Conditioned media is harvested and frozen at -20°C until the protein is to be purified.

Conditioned media is purified by affinity chromatography as described below. The media is thawed and then passed through a 0.2 µm filter. A Protein G column is equilibrated with PBS at pH 7.0, and then loaded with the filtered media. The column is washed with PBS until the absorbance at A<sub>280</sub> reaches a baseline. LGR8 polypeptide is eluted from the column with 0.1 M Glycine-HCl at pH 2.7 and immediately neutralized with 1 M Tris-HCl at pH 8.5. Fractions containing LGR8 polypeptide are pooled, dialyzed in PBS, and stored at -70°C.

For Factor Xa cleavage of the human LGR8 polypeptide-Fc fusion polypeptide, affinity chromatography-purified protein is dialyzed in 50 mM Tris-HCl, 100 mM NaCl, 2 mM CaCl<sub>2</sub> at pH 8.0. The restriction protease Factor Xa is added to the dialyzed protein at 1/100 (w/w) and the sample digested overnight at room temperature.

#### Example 4: Production of Anti-LGR8 Polypeptide Antibodies

Antibodies to LGR8 polypeptides may be obtained by immunization with purified protein or with LGR8 peptides produced by biological or chemical synthesis. Suitable procedures for generating antibodies include those described in Hudson and Bay, *Practical Immunology* (2nd ed., Blackwell Scientific Publications).

In one procedure for the production of antibodies, animals (typically mice or rabbits) are injected with an LGR8 antigen (such as an LGR8 polypeptide), and those with sufficient serum titer levels as determined by ELISA are selected for hybridoma production. Spleens of immunized animals are collected and prepared as single cell suspensions from which splenocytes are recovered. The splenocytes are fused to mouse myeloma cells (such as Sp2/0-Ag14 cells), are first incubated in DMEM with 200 U/mL penicillin, 200 µg/mL streptomycin sulfate, and 4 mM glutamine, and are then incubated in HAT selection medium (hypoxanthine, aminopterin, and thymidine). After selection,

the tissue culture supernatants are taken from each fusion well and tested for anti-LGR8 antibody production by ELISA.

Alternative procedures for obtaining anti-LGR8 antibodies may also be employed, such as the immunization of transgenic mice harboring human Ig loci for production of human antibodies, and the screening of synthetic antibody libraries, such as those generated by mutagenesis of an antibody variable domain.

#### Example 5: Expression of LGR8 Polypeptide in Transgenic Mice

To assess the biological activity of LGR8 polypeptide, a construct encoding an LGR8 polypeptide/Fc fusion protein under the control of a liver specific ApoE promoter is prepared. The delivery of this construct is expected to cause pathological changes that are informative as to the function of LGR8 polypeptide. Similarly, a construct containing the full-length LGR8 polypeptide under the control of the beta actin promoter is prepared. The delivery of this construct is expected to result in ubiquitous expression.

To generate these constructs, PCR is used to amplify template DNA sequences encoding an LGR8 polypeptide using primers that correspond to the 5' and 3' ends of the desired sequence and which incorporate restriction enzyme sites to permit insertion of the amplified product into an expression vector. Following amplification, PCR products are gel purified, digested with the appropriate restriction enzymes, and ligated into an expression vector using standard recombinant DNA techniques. For example, amplified LGR8 polypeptide sequences can be cloned into an expression vector under the control of the human  $\beta$ -actin promoter as described by Graham *et al.*, 1997, *Nature Genetics*, 17:272-74 and Ray *et al.*, 1991, *Genes Dev.* 5:2265-73.

Following ligation, reaction mixtures are used to transform an *E. coli* host strain by electroporation and transformants are selected for drug resistance. Plasmid DNA from selected colonies is isolated and subjected to DNA sequencing to confirm the presence of an appropriate insert and absence of mutation. The LGR8 polypeptide expression vector is purified through two rounds of CsCl density gradient centrifugation, cleaved with a suitable restriction enzyme, and the linearized fragment containing the LGR8 polypeptide transgene is purified by gel electrophoresis. The purified fragment is resuspended in 5 mM Tris, pH 7.4, and 0.2 mM EDTA at a concentration of 2 mg/mL.

Single-cell embryos from BDF1 x BDF1 bred mice are injected as described (PCT Pub. No. WO 97/23614). Embryos are cultured overnight in a CO<sub>2</sub> incubator and 15-20 two-cell embryos are transferred to the oviducts of a pseudopregnant CD1 female mice. Offspring obtained from the implantation of microinjected embryos are screened by PCR amplification of the integrated transgene in genomic DNA samples as follows. Ear pieces are digested in 20 mL ear buffer (20 mM Tris, pH 8.0, 10 mM EDTA, 0.5% SDS, and 500 mg/mL proteinase K) at 55°C overnight. The sample is then diluted with 200 mL of TE, and 2 mL of the ear sample is used in a PCR reaction using appropriate primers.

At 8 weeks of age, transgenic founder animals and control animals are sacrificed for necropsy and pathological analysis. Portions of spleen are removed and total cellular RNA isolated from the spleens using the Total RNA Extraction Kit (Qiagen) and transgene expression determined by RT-PCR. RNA recovered from spleens is converted to cDNA using the SuperScript<sup>TM</sup> Preamplification System (Gibco-BRL) as follows. A suitable primer, located in the expression vector sequence and 3' to the LGR8 polypeptide transgene, is used to prime cDNA synthesis from the transgene transcripts. Ten mg of total spleen RNA from transgenic founders and controls is incubated with 1 mM of primer for 10 minutes at 70°C and placed on ice. The reaction is then supplemented with 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.5 mM MgCl<sub>2</sub>, 10 mM of each dNTP, 0.1 mM DTT, and 200 U of SuperScript II reverse transcriptase. Following incubation for 50 minutes at 42°C, the reaction is stopped by heating for 15 minutes at 72°C and digested with 2U of RNase H for 20 minutes at 37°C. Samples are then amplified by PCR using primers specific for LGR8 polypeptide.

#### Example 6: Biological Activity of LGR8 Polypeptide in Transgenic Mice

Prior to euthanasia, transgenic animals are weighed, anesthetized by isoflurane and blood drawn by cardiac puncture. The samples are subjected to hematology and serum chemistry analysis. Radiography is performed after terminal exsanguination. Upon gross dissection, major visceral organs are subject to weight analysis.

Following gross dissection, tissues (*i.e.*, liver, spleen, pancreas, stomach, the entire gastrointestinal tract, kidney, reproductive organs, skin and mammary glands,

bone, brain, heart, lung, thymus, trachea, esophagus, thyroid, adrenals, urinary bladder, lymph nodes and skeletal muscle) are removed and fixed in 10% buffered Zn-Formalin for histological examination. After fixation, the tissues are processed into paraffin blocks, and 3 mm sections are obtained. All sections are stained with hematoxylin and exosin, and are then subjected to histological analysis.

The spleen, lymph node, and Peyer's patches of both the transgenic and the control mice are subjected to immunohistology analysis with B cell and T cell specific antibodies as follows. The formalin fixed paraffin embedded sections are deparaffinized and hydrated in deionized water. The sections are quenched with 3% hydrogen peroxide, blocked with Protein Block (Lipshaw, Pittsburgh, PA), and incubated in rat monoclonal anti-mouse B220 and CD3 (Harlan, Indianapolis, IN). Antibody binding is detected by biotinylated rabbit anti-rat immunoglobulins and peroxidase conjugated streptavidin (BioGenex, San Ramon, CA) with DAB as a chromagen (BioTek, Santa Barbara, CA). Sections are counterstained with hematoxylin.

After necropsy, MLN and sections of spleen and thymus from transgenic animals and control littermates are removed. Single cell suspensions are prepared by gently grinding the tissues with the flat end of a syringe against the bottom of a 100 mm nylon cell strainer (Becton Dickinson, Franklin Lakes, NJ). Cells are washed twice, counted, and approximately  $1 \times 10^6$  cells from each tissue are then incubated for 10 minutes with 0.5  $\mu$ g CD16/32(Fc $\gamma$ III/II) Fc block in a 20  $\mu$ L volume. Samples are then stained for 30 minutes at 2-8°C in a 100  $\mu$ L volume of PBS (lacking Ca<sup>+</sup> and Mg<sup>+</sup>), 0.1% bovine serum albumin, and 0.01% sodium azide with 0.5  $\mu$ g antibody of FITC or PE-conjugated monoclonal antibodies against CD90.2 (Thy-1.2), CD45R (B220), CD11b(Mac-1), Gr-1, CD4, or CD8 (PharMingen, San Diego, CA). Following antibody binding, the cells are washed and then analyzed by flow cytometry on a FACScan (Becton Dickinson).

While the present invention has been described in terms of the preferred embodiments, it is understood that variations and modifications will occur to those skilled in the art. Therefore, it is intended that the appended claims cover all such equivalent variations that come within the scope of the invention as claimed.